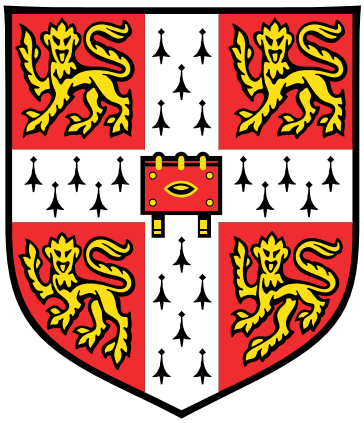
Characterisation of the structure and function of the *Salmonella* flagellar export gate protein, FlhB.



Paul Michael Bergen

Department of Pathology

University of Cambridge

This dissertation is submitted for the degree of *Doctor of Philosophy*.

Churchill College September 2017

This dissertation is dedicated to my parents, Ann and Michael, my sisters, Lisa and Diana, and my brother, Charles, for their continued love, support, and faith in all I do.

Declaration of Originality

This dissertation is the result of my own work except where reference is made to the work of others or done in collaboration with others. It is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the preface and specified in the text. I further state that no substantial part of my dissertation has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University of similar institution except as declared in the preface and specified in the text. It does not exceed the prescribed word limit for the relevant Degree Committee.

Paul M Bergen

Sep 2017

Acknowledgements

There is an old saying that it takes a village to raise a child. Well, it certainly takes a village to complete a PhD. There are many more people to thank than those in these pages, but I hope I made a start.

I would first like to thank my supervisor, Dr. Gillian Fraser, for her constant guidance, encouragement, and support during my PhD and in writing this dissertation. I would also like to thank Dr. Barry Kingston, Dr. Jim Ajioka, and Dr. Orr Yarkoni for their useful comments and help during my doctoral studies.

I have had the incredible opportunity to work with some amazing scientists during my time at Cambridge. I thank Dr. Daniel Nietlispach for his help and advice with the collection and analysis of NMR data and Dr. Lewis Evans and Dr. Nick Greene for helping me lay a strong foundation at the beginning of my PhD. I would also like to thank my colleagues Dr. Katie Kemplen and Owain Bryant for their invaluable assistance and encouragement. My science was made stronger because of them.

I would not have stayed motivated or sane during my PhD without Eliza, Angela, Andrew, Alexis, Jacqui, Rebekah, Derrick, Talia, Mohammad, Musa, Katie, Michelle, Callie, and the “Hillbillies”. For all those discussions (six seconds until they become academic), random trips to Morocco, Vienna, and places afar, boozy brunches, and so much more, I owe you all my eternal gratitude.

Finally, I wish to thank my scholarship, Gates Cambridge, for supporting my doctoral studies and for providing a vibrant, international community of excellent Scholars who take our shared commitment to improving the lives of others to heart. The Gates Cambridge staff and all the Scholars I served with on the Scholars’ Council are some of the most passionate and dedicated individuals I know. My international education experience was world-class because of their tireless work and, I believe, Cambridge is a better place because of their contributions.

Summary

Flagella, the helical propellers that extend from the bacterial cell surface, illustrate how complex nanomachines assemble outside the cell. The sequential construction of the flagellar rod, hook, and filament requires export of thousands of structural subunits across the cell membrane and this is achieved by a specialised flagellar Type III Secretion System (fT3SS) located at the base of each flagellum. The fT3SS imposes a crude ordering of subunits, with filament subunits only exported once the rod and hook are complete. This “export specificity switch” is controlled by the FlhB component of the fT3SS export gate in response to a signal from the exported molecular ruler FliK, which monitors the length of the growing hook. This study seeks to clarify how rod and hook subunits interact with FlhB, and how FlhB switches export specificity.

Rod and hook subunits possess a conserved gate recognition motif (GRM; Fxxxφ, with φ being any hydrophobic residue) that is proposed to bind a surface-exposed hydrophobic patch on the FlhB cytosolic domain. Mutation of the GRM phenylalanine and the final hydrophobic residue resulted in impaired subunit export and decreased cell motility. Isothermal titration calorimetry was performed to assess whether subunit export order is imposed at FlhB. These experiments showed that rod and hook subunits bind to FlhB with micromolar dissociation constants (5-45 μM), suggesting transient interactions. There was no clear correlation between subunit affinity for FlhB and the order of subunit assembly in the nascent flagellum. Solution-state nuclear magnetic resonance (NMR) spectroscopy supported prior data showing that rod and hook subunits interact with FlhB’s surface-exposed hydrophobic patch. NMR also indicated that residues away from the patch undergo a conformational change on subunit binding.

FlhB autocleaves rapidly in its cytosolic domain, and the resulting polypeptides ($FlhB_{CN}$ and $FlhB_{CC}$) are held together by non-covalent interactions between β-strands that encompass the autocleavage site. The autocleavage event is a prerequisite for the export specificity switch, but its function is unclear. Analysis of the cellular localization of $FlhB_{CN}$ and $FlhB_{CC}$ revealed that $FlhB_{CC}$ dissociated from the membrane export machinery, but only in the presence of FliK. Biochemical and biophysical

studies of FlhB variants that undergo export specificity switching in the absence of FliK showed that these FlhB “autonomous switchers” were less stable than wildtype FlhB and their $FlhB_{CC}$ domain could dissociate from the export machinery in the absence of FliK. The results suggest that the export specificity switch involves a FliK-dependent loss of $FlhB_{CC}$ from the export machinery, eliminating the binding site for rod and hook subunits.

Table of Contents

List of Abbreviations Commonly Used

^{1}H	Hydrogen-1 (proton)
^{13}C	Carbon-13 isotope
^{15}N	Nitrogen-15 isotope
CD	Circular dichroism
C-terminus	Carboxyl-terminus
CV	Column volume
DNA	Deoxyribonucleic acid
D_2O	Deuterium oxide
$FlhB_C$	*Salmonella* export gate FlhB cytoplasmic domain residues 211-383
$FlhB_{CN}$	*Salmonella* export gate FlhB cytoplasmic N-terminal domain residues 211-269
$FlhB_{CC}$	*Salmonella* export gate FlhB cytoplasmic C-terminal domain residues 270-383
fT3SS	flagellar Type III secretion system
GRM	Gate Recognition Motif
GST	Glutathione S-transferase
HBB	Hook-Basal Body
HMQC	Heteronuclear multiple-quantum correlation spectroscopy
SOFAST-HMQC	2D Selective Optimized-Flip-Angle Short-Transient Heteronuclear Multiple Quantum Correlation
HNCA	Heteronuclear Carbon Alpha 3D spectroscopy
HN(CA)CB	Heteronuclear Carbon Alpha/Carbon Beta 3D spectroscopy
HN(CO)CA	Heteronuclear Carbonyl-Carbon Alpha 3D spectroscopy
HN(COCA)CB	Heteronuclear Carbonyl-Carbon Alpha/Carbon Beta 3D spectroscopy
ITC	Isothermal Titration Calorimetry
K_d	Dissociation constant
kDa	kilo Daltons
Ni-NTA	Nickel Nitrilotriacetic acid resin
NMR	Nuclear Magnetic Resonance
NOESY	Nuclear Overhauser effect spectroscopy
NPTH	Asparagine-Proline-Threonine-Histidine amino acid motif
N-terminus	Amino-terminus
A_{600}	Absorbance at 600 nm
PBS	Phosphate-buffered saline
PBST	Phosphate-buffered saline with 0.05% Triton X-100
PCR	Polymerase Chain Reaction

PMF	Proton Motive Force
RBS	Ribosome binding site, Shine-Dalgarno sequence
SDS PAGE	Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis
SOC	Super Optimal broth with Catabolite repression
T3SS	Type III secretion system
T3S4	Type III secretion substrate specificity switch
TCA	Trichloroacetic acid

List of Figures

List of Tables

Chapter 1

Introduction

1.1 The bacterial flagellum

Motility is a common feature of life. The ability of organisms to swim, to fly, or to walk to new, more favourable environments presents them with an advantage over their non-motile competitors. They can also avoid harmful chemicals like antibiotics, or evade predators, things their non-motile competitors must evolve defenses to combat. Organisms across the domains of life have evolved many features that impart motility or allow them to be moved by external physical forces.

Bacteria have evolved long, self-assembling helical propellers called flagella to move through their environment. These elegant structures span the cell envelope and extend far into the environment, requiring the orchestration of dozens of genes and proteins to assemble the three substructures of the flagellum – the basal body, the hook, and the filament. Thousands of subunits must be exported from the cell to construct a single flagellum since bacterial flagella are assembled at the distal end of the growing structure. This poses a grand challenge, made even larger in bacteria possessing multiple flagella. How does a cell express thousands of protein subunits and energise their export? Just as important, how does a cell coordinate the export of so many proteins so that a competent structure is assembled?

Long championed as a representation of irreducible complexity and the hand of a creator, the bacterial flagellum is an example of evolution at work over millions of years. Since its discovery in the late 19th century, the flagellum has been studied by bacteriologists, biophysicists, and others to demystify its complexity. Subunits vary across species, some lost, duplicated, or replaced as needed. Rather than providing proof of the hand of a creator, studying the bacterial flagellum has given us a strong example of natural selection leading to a diversity of forms while maintaining a unified function (1). Over 25 proteins must come together in a web of interactions to assemble and operate the final product. What sort of interactions must occur and how do these interactions produce a functional flagellum?

1.2 Flagellar motility and arrangement

1.2.1 Rotary works

There are two types of flagellar motility – swarming and swimming. Swimming describes the movement through liquid environments, while swarming is the ability of bacteria to move across surfaces. This study investigated swimming motility, so that will be the focus of the following section.

Flagellar swimming motility is often described as “tumble and run” (2). During run behavior, the flagellum can rotate clockwise (CW) or counterclockwise (CCW), converting rotary motion into thrust and leading to smooth swimming (2, 3). When the motor stops and begins to switch direction or slows down, the cell tumbles and ceases movement. Once the motor restarts and enough thrust is generated, the cell can return to a “run” and swim smoothly towards a chemotactic signal (2, 4, 5).

Far from a random process, motility is guided by a response to chemotactic signals such as pH, temperature, and molecules sensed by methyl-accepting chemotaxis proteins (MCPs) (2, 6). Chemotaxis and rotational direction of the flagellar motor is controlled by the CheA-CheY two component system (7). MCPs modulate the autophosphorylation activity of CheA depending on whether they have sensed an attractant or repellent. Upon a MCP sensing a repellent, CheA undergoes autophosphorylation before transferring the phosphate group to CheY. Phosphorylated CheY then interacts with the flagellar motor and “switch” proteins, FliG/FliM/FliN. Phosphorylated CheY interacts directly with FliM, probably leading to a change in the interaction of FliM with the motor protein, FliG (2, 8). This cascades to alter the interaction between FliG and the stator subunit MotA, eventually altering the stator-motor complex to switch direction from CCW to CW rotation. If the MCP senses an attractant signal, CheA autophosphorylation is repressed. CheA cannot then phosphorylate CheY and the motor will not switch directions, allowing the cell to continue to move toward the attractant (2).

Flagellar motility is driven by the proton motive force (PMF) via the MotA and MotB stator complex or the sodium motive force (SMF) via the PomA and PomB stator complex (9–11). Not all bacteria produce flagella with a single ion motor. Sodium and

proton motive force flagella are simultaneously present in some marine bacteria. To slow down or switch direction in these cases, the bacterium can slow its proton-drive motor or switch the direction of its sodium-driven motor (12). Although their energy sources are different, functional chimeras containing parts of both stator complexes can also be engineered (13). This suggests the drive mechanism is shared amongst the various ion-driven motors.

Each unit of the PMF motor complex is formed of a 2:1 ratio of MotA to MotB, with the most probable stoichiometry being $MotA_4MotB_2$ (14). At most, 11 complexes of MotA/MotB assemble at the base of each flagellum (15). The individual units in the complex are not fixed to the motor, but rather transiently engage with the motor with an average dwell time of 30 s. Any of the 22 units of MotB at the motor can be exchanged with a membrane pool of approximately 200 unbound stators, for example (16). This freely diffusing membrane pool of stators might be an insurance policy of sorts and suggests the MotA/MotB motor complex is not as static as previously claimed (16). As the stator complex in the flagellar motor wears down, its components may be exchanged for fresh ones from the membrane pool to continue the flagellum's smooth, efficient operation. The cell can then continue to move through its environment with little mechanical impediment.

Both stator proteins are inserted into the cell membrane and contain conserved, charged residues necessary for torque generation and motor rotation (2). In MotA/PomA, these charged residues are located on a cytoplasmic loop between the second and third transmembrane domains where they can interact with the motor protein, FliG (17–19). MotB/PomB only contains one transmembrane domain where a conserved aspartate is required for proton or sodium translocation across the cell membrane (20–23). The C-terminus of MotB/PomB sticks into the periplasm and contains a peptidoglycan-binding motif to anchor the stator complex (24, 25). Together, these stator complexes form the ion channels that drive rotation by harnessing the energy provided by protons or sodium ions moving across the cell membrane.

The electrostatic interactions between the motor protein, FliG, and the stator complex ultimately drive the rotation of the flagellum. The full picture of torque generation by the flagellar motor is still unclear, but there is some information on the speed at which the motor can operate and the amount of torque generated. A low estimate of the proton-driven flagellar motor suggests it can rotate at approximately 63 Hz, generating 1300 pN·nm of force, although higher estimates of 1.5 x 10^5 pN·nm s^{-1} of power have also been observed (15, 26). Speeds up to 1700 Hz have been recorded for the sodium-driven motor (27). These rotary speeds translate into a range swimming speeds. Peritrichous organisms such as *Salmonella*, which utilise the PMF, have been measured swimming at 55 μm s^{-1} at 35°C, or over ten-times their body length per second (28). This is a casual jog compared to organisms with single, polar flagella powered by the SMF such as *Vibrio alginolyticus*, which has been clocked at 116 μm s^{-1} at 35°C (29).

1.2.2 Flagellar arrangement

The number and arrangement of flagella varies greatly by cell and by species. It isn't clear why some species assemble more than one flagellum when a singular flagellum is sufficient to swim through a liquid environment (30, 31). Although this topic isn't as well studied as flagellar subunit export and assembly, recent research has begun to explore the mechanisms that determine flagella positioning in the cell.

As the flagellum structure crosses the cell envelope, flagella are immobilised in the cell wall. Once assembly has proceeded through the cell wall, the position of that particular flagellum has been settled, but the exact mechanism behind this choice is largely unknown. Broadly, flagellar arrangement can be classed in four groups. Flagella can be singular and polar or singular and situated along the length of the cell. They can also be multiple and polar, called lophotrichous, or multiple and spread across the cell, called peritrichous. The biological importance of flagellar patterning and arrangement is poorly understood, but we do know that positioning and number of flagella is not a random process (31).

This patterning is under the control of two subunits, with slight variations made for the arrangements found across bacteria. FlhF and FlhG have been shown to dictate basal

body positioning. FlhF is a GTPase and FlhG serves as its ATPase antagonist (31–33). Although much of their function remains unknown, it's clear that FlhF and FlhG co-localise at the base of the flagellum (34, 35).

How may the FlhF-FlhG pairing regulate flagellar patterning and number? In species with polar flagella, the loss of FlhF causes flagella to be inserted along the length of the cell rather than at the poles (33, 34, 36). The loss of FlhG increases the number of assembled flagella in species with a single flagellum (32, 35, 36). In peritrichous bacteria, however, the loss of FlhF leads to polar flagellar and the loss of FlhG contributes to the aggregation of basal bodies along the cell (31). In addition to their role in determining flagellar placement and number, FlhF and FlhG may serve to determine the number of flagella daughter cells receive after binary fission is complete, therefore determining the motility of individual cells, although not necessary the motility of the population (31, 35). So, although the FlhF-FlhG relationship and control is maintained across species, it is fluid enough to allow for the range of flagellar arrangement seen in bacteria (31).

Bacteria have evolved a variety of arrangements, numbers, and modifications to their flagella. Flagella, in any arrangement or number, can generate substantial thrust and propel the organism into new frontiers. These frontiers often include the guts of mammals, where flagella play important roles in pathogenesis.

1.3 Motility in pathogenesis

Pathogens that can swim have an inherent advantage over their non-motile counterparts; namely, they do not have to rely on Brownian motion or mechanical forces to reach a potential site of infection (37). The same chemotaxis system that leads bacteria to a favourable environment is also used by pathogens to sense their preferred niche. *Helicobacter pylori* can detect sites of stomach injury to establish infection and *Campylobacter jejuni* detects mucins and glycoproteins to find the intestinal mucus-filled crypts it preferentially colonises (38, 39). *Salmonella enterica* serovar Typhimurium requires flagella with an active chemotaxis system to swim through the mucosal layer to reach the murine intestinal epithelium, allowing it to contact host cells and begin to inject effector molecules (40).

The flagellum is not simply employed to move from point A to point B. It plays a complex role in the pathogenesis of many bacteria including *S. enterica*, *C. jejuni*, *Pseudomonas aeruginosa*, and *Escherichia coli* (41). Beyond swimming, flagella have been shown to be important for adhesion, biofilm formation, effector molecule secretion, and modulating the immune response (37).

The filament, the largest of the flagellar substructures, is important for adhesion to host cell membranes during infection (37). In Enteropathogenic *E. coli* (EPEC), flagella are highly induced upon interaction with epithelial cells to adhere the pathogen to the host cell. Girón and colleagues noticed that the number of flagella was reduced when EPEC was cultured with methanol-fixed cells, suggesting that some soluble molecule secreted by living cells is responsible for the increase in flagella number. They further went on to show that motility medium preconditioned by the growth of HeLa cells could induce motility in a strain previously unable to swim in the motility medium (42). Enterotoxigenic *E. coli* (ETEC) secretes EtpA, an adhesion exoprotein, that interacts with flagellin, the filament subunit. This interaction between the two ETEC proteins is crucial for adherence to the host cell and eventual colonisation (43). Adherence by flagella to the host cell has also been identified in other pathogens such as *P. aeruginosa* and *C. jejuni* (44, 45). Recently, work by E. Wolfson and colleagues suggests that the filaments of Enterohaemorrhagic *E. coli* (EHEC) can interact with components of the host cytoskeleton and might even penetrate host cell membranes, although it is unclear by what mechanism (46). Once contact is established, pathogens like *Salmonella* and *Shigella* can induce the cell to engulf them to continue with infection and either stay hidden in specialized vesicles or move into the cytosol for growth and replication (47).

Of course, reaching the site of infection or invading a host cell is only the beginning of infection. To survive, the pathogen must evade the immune system. Having a long filament made up of many repetitive regions invites immune detection and response, stymying infection. Motile pathogens have evolved a few strategies to avoid death by host immunity (37). A common strategy pathogens employ is the downregulation of flagellar gene expression. *Salmonella* Typhimurium greatly downregulates flagellar and chemotaxis gene expression once inside macrophages (48). EHEC

downregulates flagellar gene expression after attachment to and formation of microcolonies on host intestinal epithelial cells (49). *Legionella pneumophila* actively senses its external temperature and begins to downregulate flagellar expression at 37°C (50). In continuing infection, these pathogens actively sense their external environment and dispense with flagella production to evade the immune system.

Pathogens have also evolved subtler means than downregulating gene expression to evade the immune system. *Salmonella* Typhimurium expresses two flagellins, FliC and FljB, with different antigenicities (51, 52). This leaves the bacterium exposed to an immune response, but it also makes it harder for the host immune system to detect and react to the pathogen (37). *Salmonella* cells are also known to exhibit bistability, or maintaining a population ratio of motile to non-motile cells depending on environmental conditions. Individual motility, then, can be sacrificed to allow the population to survive (53). Another method pathogens have evolved to survive in a host is to alter their flagellin sequence to avoid detection by host toll-like receptor 5 (TLR5) (54). Although mutations present a risk to the cell's survival, natural selection may well favour a pathogen that can avoid immune detection while maintaining its ability to swim and colonise its host.

Clearly, flagella are vital to the pathogenesis of many bacteria. The following sections provide further detail into *Salmonella enterica* serovar Typhimurium, its flagella, and the complex molecular interactions that are required for the cell to produce a competent flagellum. Comparisons will continue to be made to other organisms, where appropriate, but as this study focuses on the *Salmonella* Typhimurium flagellum, it will be the focus of the remaining sections.

1.4 Salmonella enterica serovar Typhimurium

550 million people, and 220 million children under the age of five, contract diarrhoeal diseases each year. About 33 million people eventually die from their disease. *Salmonella* is one of the leading causes of these diarrhoeal diseases globally, affecting those in developed and developing countries. *Salmonella* is genus of common Gram-negative, motile, rod-shaped pathogens. They readily survive in the soil or water for weeks and can colonise different mammalian species, allowing them

to spread from animals to humans. Transfer of *Salmonella* is often due to poor hygiene after direct contact with infected individuals or food.

Salmonella enterica serovar Typhimurium, like most *Salmonella* serovars, mostly causes sporadic diarrhoeal disease in humans that is faeco-orally spread, usually self-limiting, and typically lasts less than a week. It can, however, cause severe disease in the young, the elderly, and those with compromised immune systems (55). Additionally, in mice, it causes a disease similar to typhoid fever in humans. This makes *Salmonella* Typhimurium a suitable model organism to study the much more severe typhoid fever without the risk of contracting the disease.

Prevention requires efforts across the food supply chain such as better hygiene, prevention of faecal contamination of meat, fruit, and vegetables, and keeping processing surfaces clean and dry. International efforts to track and reduce *Salmonella* outbreaks are ongoing. They include giving national health organisations the tools to monitor food processing, educating consumers and industry on better hygienic procedures, and establishing improved safety standards (55).

The most common treatment of salmonellosis is rehydration and, if the patient experiences severe dehydration, electrolyte replacement. Antibiotics are only recommended for those patients with severe disease or in whom the bacterium has spread from their gut to the blood stream. This is because of the growing concern of antibiotic resistance and the self-limiting nature of most salmonellosis (55). According to the National Antimicrobial Resistance Monitoring System in the United States, *Salmonella* Typhimurium is more likely to be resistant to antibiotics and to possess resistance against more classes of antibiotics than other *Salmonella* serovars (56). Data varies from nation to nation and year to year, but *Salmonella* Typhimurium is either the most common or second-most common serovar of *Salmonella enterica* to cause foodborne disease in humans (55, 57). This ubiquity only makes the spread of antibiotic resistance among *Salmonella* Typhimurium more troubling. There is a clear need for better surveillance and faster treatment of humans and animals with salmonellosis coupled with a greater understanding of *Salmonella* Typhimurium biology.

The flagellum is an important component of *Salmonella* Typhimurium pathogenesis. By understanding its structure, and how the cell goes from flagellar gene expression to subunit assembly, we will be better equipped to develop novel, rational therapies to combat this global menace. The strain used in this study as a wild type motility model, SJW1103, is a *Salmonella* Typhimurium strain originally described in 1984. It is phase-locked, meaning it only expresses FliC flagellin and assembles a homogenous filament (58).

1.5 Bacterial secretion systems

Gram-negative bacteria have evolved a multitude of highly specialised systems to secrete substrates into the extracellular environment or a host cell (Figure 1.1). These substrates, which can include small molecules like antibiotics or large proteins, are typically secreted as a response to some external stimuli. Secreted substrates may lead to the assembly of an intricate nanomachine, like the flagellum or a pilus, or could remove a harmful molecule from the cytoplasm. They can also cause host cells to react in such a manner as to initiate or prolong infection (59). Current work has identified over a dozen secretion systems in Gram-negative bacteria, although more certainly remain to be identified (60). The diversity of secretion systems give Gram-negative bacteria a veritable arsenal of options to regulate their internal and external environments.

Secretion systems can be broadly described with two mechanisms. In one-step secretion, substrates are secreted directly from the bacterial cytoplasm to outside the cell. Type I, III, IV, and VI secretion systems all work in this manner (59). Other secretion systems, like Type II and V, use a two-step mechanism. In this case, substrates are first secreted into the periplasmic space before being secreted outside the cell. These systems require an inner membrane transporter to secrete substrates into the periplasm or insert them into the inner membrane (61). The SecYEG translocon is often used to transport unfolded polypeptides across the inner membrane, while the twin-arginine translocation (Tat) system is employed to transport folded proteins across the inner membrane (59, 62).

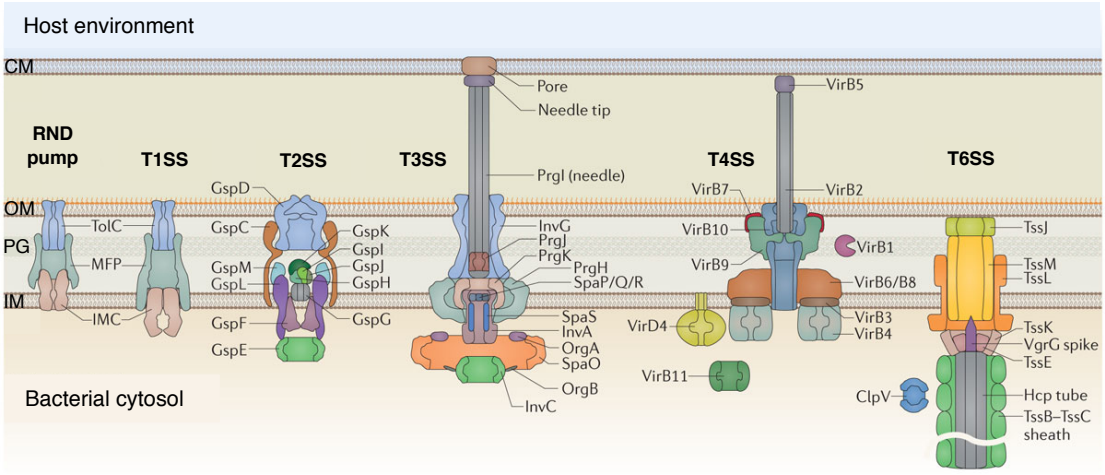


Figure 1.1 Bacteria have evolved many systems to export molecules

Different secretion systems serve many different functions. Some are used to remove harmful molecules, such as antibiotics, from the cell, whilst others are used to transport effector molecules from the bacterial cell into the host cell in order to cause infection. Others secrete the subunits necessary to assemble intricate nanomachines for use in adhesion, motility, and more. Secretion systems can also be used to share genetic material, spreading genes useful in host colonisation, nutrient utilisation, and antimicrobial resistance. Some notable systems missing from this figure are the Sec system, flagellar type III secretion system (shown in figure 1.2A), type VII secretion system, and the newly characterised type IX secretion system found in the *Bacteroides* phylum.

RND pump is the resistance-nodulation-division pumps; *T1SS* is the type I secretion system; *T2SS* is the type II secretion system; *T3SS* is the injectisome type III secretion system; *T4SS* is the type IV secretion system; and *T6SS* is the type VI secretion system.

CM is the host cell membrane; *OM* is the bacterial outer membrane; *PG* is the peptidoglycan cell wall; and *IM* is the bacterial inner membrane.

Figure adapted from Costa et al., 2015.

Energy sources for substrate secretion vary by system. Some, like type IV secretion systems, energise secretion by ATP hydrolysis at dedicated ATPases (63). Others, like RND pumps, use proton gradients to power secretion (64). And in other cases, like type V secretion systems, energy is provided by the protein substrate folding in the extracellular environment, leading these systems to be termed “autotransporters” (65).

In many cases, secretion systems must adopt different conformations to control substrate secretion. The TolC outer membrane channel of Type I secretion and RND tripartite pumps, for example, assumes an open conformation when in complex with the inner membrane component and periplasmic adaptor proteins. This complex, creating a channel that spans the cell envelope, is only able to form upon substrate binding (59). In the absence of this complex, the TolC channel is closed at the periplasmic side to prevent substrates from inadvertently entering the channel (66).

Proteins pose a challenge that small molecules do not. Many of these secretion systems have very narrow central channels, restricting the size of molecules that can pass through. The injectisome type III secretion system in *Salmonella* has an inner channel diameter of ~25 Å, much too small for folded protein to pass through (67, 68). Some secretion systems are capable of secreting folded protein; however, they usually begin secretion with unfolded polypeptides. The usher of the chaperone-usher pathway used to construct type I pili can mediate secretion of folded protein. In this pathway, unfolded polypeptides are first secreted by SecYEG into the periplasm (69). There, a pilus chaperone folds the subunit, thereby stabilising it against periplasmic proteases, and delivers it to the usher for secretion and polymerisation into the growing pilus (70).

Bacterial flagella are assembled at the distal end, necessitating the secretion of thousands of subunits from the cell. Flagella use dedicated type III secretion systems (T3SS) to overcome this challenge. The following section discusses type III secretion in greater detail and how these systems manage to export proteins up to 20 μm away from the cell.

1.5.1 Type III secretion systems

One of the most common secretion systems found in bacteria is the type III secretion system. There are two evolutionary-related T3SSs which secrete subunits across the cell envelope for the assembly of two nanomachines – the flagellum and the injectisome.

The injectisome is used by many Gram-negative pathogens to inject effector molecules across the cell membranes of eukaryotic host cells (71). These injectisomes are major players in the infection of animal and plant cells. The injectisome is made up of the core needle structure, a needle tip assembly for the construction of the translocation pore, and an assortment of secreted effector molecules (72). The range of effector molecules and their targets, often implicated in disease but also in symbiosis between bacterium and host, varies greatly across bacterial species (73). These effector molecules transit the central channel of the needle structure and enter the host cell through a pore formed by translocon subunits once contact with the host cell has been made (74). Once inside the host cell, these effector molecules are free to bind to their targets to elicit a favourable response from the host cell (71).

Many proteins necessary for type III secretion are found in both systems, sharing structural and functional similarities, although not always significant sequence homology (Figure 1.2A). The flagellar and injectisome basal bodies are surrounded by ring structures that are embedded in the bacterial cell envelope (74). The export machinery of both systems contains a dedicated ATPase to separate subunits from their chaperones and begin to unfold subunits prior to export (75, 76). Export gates serve to coordinate export by recognising subunits via N-terminal secretion signals or chaperone-subunit complexes and directing the subunits to the export channel (76–89). Both systems use the proton motive force to begin to translocate subunits across the cell envelope, either to be assembled into the nascent structure or to be secreted into the environment (90–92).

The evolution of the two T3SSs has been hotly debated for some time. It was once thought that the two systems shared a common ancestor or even that the injectisome came first, but recent work advocates for a flagellum-first hypothesis. Flagellar genes

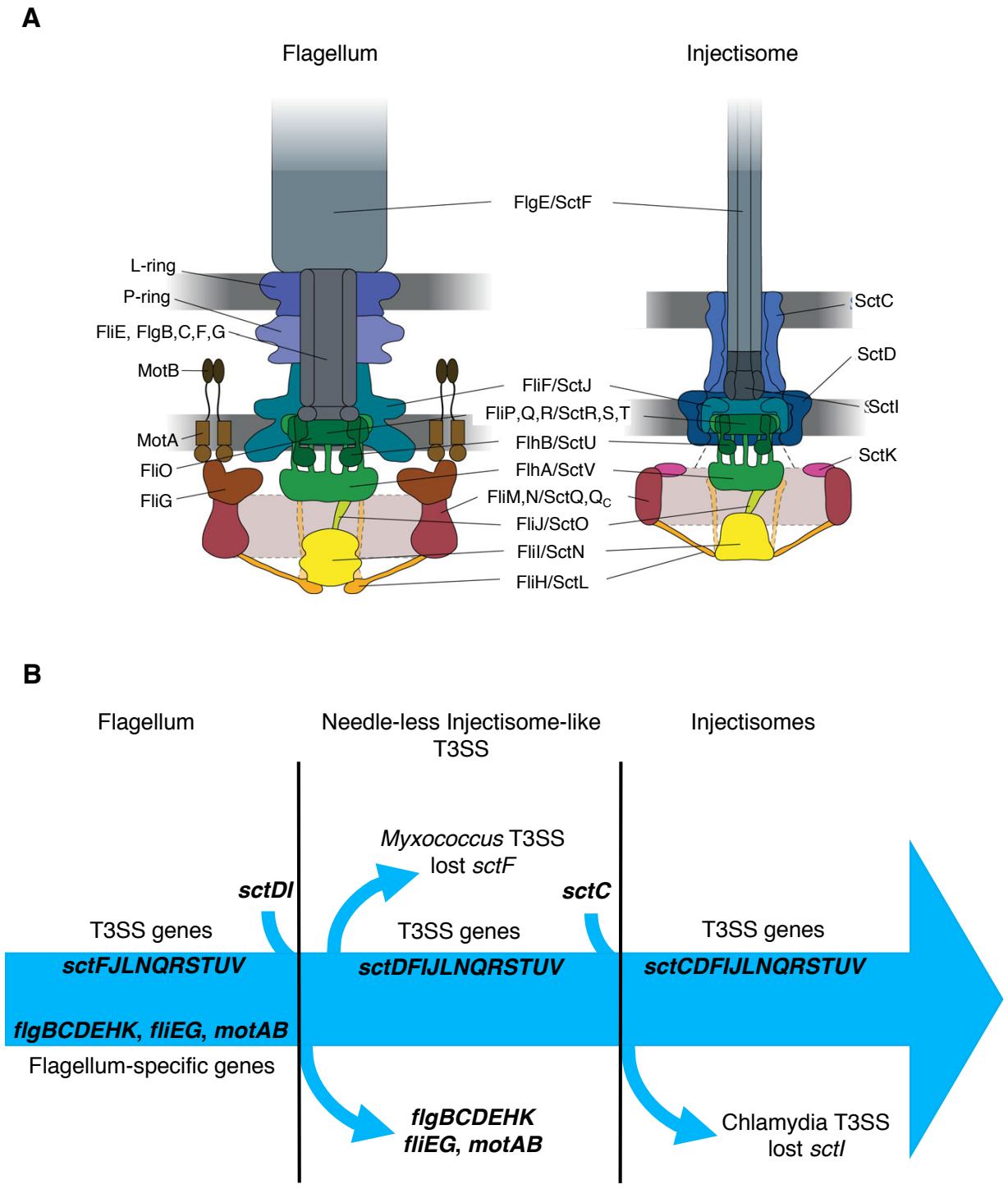


Figure 1.2 Type III secretion systems (T3SSs) are closely related

(A) Comparison between the flagellum and injectisome basal structures. Subunits found in both systems are listed in the middle of the two structures. The unified *sct* nomenclature for the injectisome is used. Adapted from Diepold and Armitage 2015.

(B) Current hypothesis of T3SS evolutionary history. SctC was acquired multiple times by different subfamilies of injectisomes. The Chlamydia subfamily is ancestral to the other animal and plant pathogen T3SSs. Adapted from Abby and Rocha 2012 and Diepold and Armitage 2015.

are commonly found on the chromosome, while injectisome genes are found on pathogenicity-islands or virulence plasmids, supporting the flagellum-first hypothesis (74). Modern large-scale genomic analysis has revealed that the flagellar T3SS is closer to the ancestral system than the injectisome T3SS (Figure 1.2B). This data uncovered that more bacteria possess flagella than possess injectisomes. Flagella are found in 14 Gram-positive and Gram-negative bacterial taxa, while injectisomes have only been found in three Gram-negative bacterial taxa (93).

This genomic analysis gave clear evidence, via the shared ATPase found in both T3SSs, that the injectisome evolved from an ancestral flagellar system in three major steps. The injectisome was first adopted as a secretion system in free living bacteria. This first adaption occurred when the injectisome ancestor lost its motility genes, retaining the core T3SS genes. Secretins were then recruited from other secretion systems, allowing the injectisome to further evolve in animal and plant pathogens for association with a range of host cells (Figure 1.2B) (93).

Because of their homology and commonality among human pathogens, the T3SS is an attractive target for designing novel therapeutics, especially in the wake of growing antimicrobial resistance and a dry antibiotic development pipeline (94). Potentially, a compound that targets one of the core T3SS subunits could cripple both the flagellum and injectisome. Pharmacological inhibitors have been discovered for some injectisomes, ranging from targeting gene regulation and transcription to components of the injectisomes themselves (95). The *Salmonella* Typhimurium flagellum was identified as a therapeutic target as early as the 1950s, with efforts concentrating on how to prevent its formation and, therefore, keep *Salmonella* populations non-motile (96). These experiments were conducted on bacterial cultures and in host models, such as mice. More recent work has shown it is possible to target the *Salmonella* flagellum using antibodies to enhance the immune response to pathogenic bacteria by phagocyte-like cells in culture (97). High-throughput screens to find new classes of flagellar inhibitors have shown some promise against the sodium motor of *Vibrio* but were less successful in finding compounds with anti-motility activity in *Salmonella* (98).

Type III secretion systems can also be hijacked as protein translocation pumps to deliver specialised protein cargo. They offer attractive opportunity for the biotechnology industry as these systems are capable of rapidly exporting a lot of protein into the environment, eliminating the need to harvest proteins from cells. The flagellar type III secretion system (fT3SS) is remarkably fast, capable of exporting up to 1,700 amino acids per second (99). For comparison, the haemolysin A type I secretion system exports roughly 17 amino acids per second and the Sec machinery is capable of exporting around 220 amino acids per second (100, 101). Protein engineering is relatively simple as well. By adding the specific N-terminal signal for fT3SS recognition and export or fusing the target protein to a secreted flagellar subunit, protein can be exported into the supernatant so long as it can be unfolded by the fT3SS for transit through the channel.

The fT3SS has been used to export peptides derived from snails, snakes, and spiders by fusing them to FlgM, a secreted subunit not assembled into the final flagellum structure (102). Dobó and colleagues were able to use the fT3SS system to secrete different proteins such as GFP and maltose-binding protein by placing a disordered N-terminal portion of flagellin (residues 26-47) at the N-terminus of fusion protein (103). Early flagellar subunits are not chaperoned and possess a five-residue recognition sequence that could be used in biotechnology (86). The injectisome has been used to deliver vaccines or cytokines into host cells and spider silk into the culture supernatant for harvesting (74). These are only a few ways in which T3SS can be targeted or hijacked for our imaginative purposes.

The complexity of type III secretion systems requires fine coordination beginning with gene expression. These systems are not constitutively expressed (59). Gene expression, therefore, is often controlled by environment signals to manage the energy cost to the cell of building these nanomachines. The next section discusses how the cell reaches the decision to begin flagellar and type III secretion gene expression and efforts to tune gene expression to various cues.

1.6 Flagellar gene arrangement and expression

Salmonella flagella require over 60 genes for proper structure and function (104). Flagellar gene expression can be broadly classed into three promoter classes that matches structural hierarchy (Figure 1.3) (105). The entire system is under control of the class I $FlhD_4C_2$ heterohexameric master regulator complex, but not individual FlhD or FlhC subunits (106).

The FlhDC complex is activated and regulated by many different cell signals based on the environment and life stage (motile vs non-motile). Examples of positive regulators of *flhDC* transcription in *Salmonella* include cAMP-CRP and Fis (107, 108). An example of a negative regulator is RflM, which acts to control the $FlhD_4C_2$ complex's own autoinhibitory effect on its transcription (109). The *Salmonella* cell cycle is also a strong transcriptional regulator, both positive and negative, of *flhDC*. Flagella production is energetically costly and is therefore reduced in favour of other cell processes during most growth phases. Flagellar gene expression is highest in mid-exponential phase (110). The $FlhD_4C_2$ complex is further regulated post-translationally by the ClpXP ATP-dependent protease to control the level of class II and III transcripts. ClpXP can recognise and regulate the heterohexameric complex, but not individual FlhD or FlhC proteins (111).

Once FlhDC is activated, it binds to the chromosome upstream of the *fliA*, *fliL*, and *flhB* operons and promotes the transcription of class II genes with σ^{70} and RNA polymerase (112, 113). This activation requires an interaction with the C-terminal domain of the RNA polymerase α subunit (113). Class II genes include those of the hook-basal body (HBB) and flagellar motor complex (105). Upon HBB completion, transcription of class III genes by the class-specific σ^{28} RNA polymerase is activated (114, 115). The shift between class II and class III gene expression is regulated by FlgM, the first of the "late" subunits (115). FlgM is only secreted once the hook has reached a mature length and the fT3SS has switched export specificity from "early" to "late" export, thereby coupling gene expression control to structure assembly (105).

FlgM serves as an anti-sigma factor to FliA (the σ^{28} factor), thus inhibiting class III gene expression by tying up FliA in the cytoplasm and actively causing it to dissociate

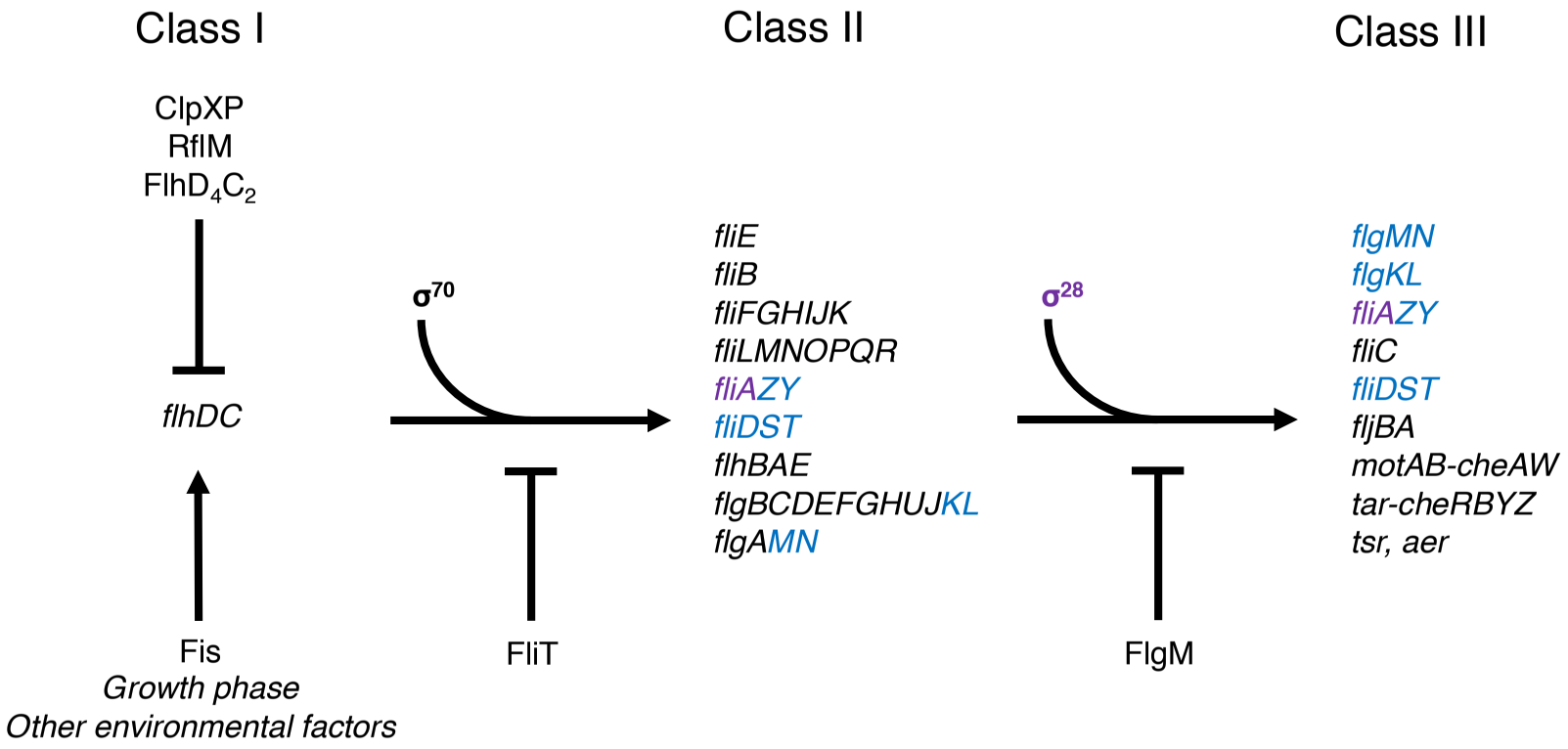


Figure 1.3 Flagellar gene regulation is complex

Genes are separated into three classes. The class I gene product, the $FlhD_4C_2$ heterocomplex, serves as the global regulator for flagellar synthesis. Once activated, $FlhD_4C_2$ can promote class II gene expression by recruiting σ^{70}. Class III genes are only expressed when σ^{28} (FliA) recruits RNA polymerase. FlgM is the anti-σ^{28} factor and blocks FliA from recruiting RNA polymerase before the completion of the hook-basal body (HBB) complex. Once the HBB is complete, FlgM is secreted and FliA can activate class III expression. Genes in blue are under the control of class II and class III promoters. σ^{28} (FliA) is coloured purple. This figure is adapted from Frye et al., 2006.

from the RNA polymerase (116–118). After the fT3SS export specificity switch and FlgM secretion, FliA is free to associate with the RNA polymerase and activate class III gene expression, resulting in flagellin expression (105). FliT, the chaperone for the filament cap FliD, also serves to regulate gene hierarchy (119). Once the fT3SS system begins to export FliD, FliT is free to interact with the $FlhD_4C_2$ complex and reduce both its ability to activate class II genes and the autoinhibition of its own transcription (105, 120).

In this way, FliT suppresses class II gene expression by inhibiting the $FlhD_4C_2$ global regulator complex. Interestingly, the repression of the autoinhibitory effect of $FlhD_4C_2$ serves to re-initiate class I flagellar expression (105). It is not fully understood why this happens, but it may be connected to the genetic regulation interplay between flagellar motility and host cell invasion by the injectisome. FliT serves to repress the expression of the *Salmonella* Pathogenicity Island 1 (SPI1) injectisome that is essential for *Salmonella* to invade host intestinal epithelium and cause disease (121). Additionally, the $FlhD_4C_2$ complex is one of many factors that activate SPI1 expression. Therefore, FliT-mediated suppression of FlhDC and SPI1 may serve to coordinate motility to first find a host cell to invade and then eventual invasion by controlling the expression of the different *Salmonella* T3SS (122).

Complicating gene expression is that some late subunits, such as FlgM, FlgK, and FlgL, are under the control of class II and class III promoters (105). But flagellin is strictly under class III control, so the filament cannot begin export and assembly until the HBB and junction proteins are assembled (104, 123, 124). This regulatory regime is made even more difficult in peritrichous bacteria like *Salmonella* where many flagella may be in different stages of assembly. How *Salmonella* regulates each individual flagellum, and the contribution of local and global control, is still being investigated.

Gene expression is only one way by which *Salmonella* cells regulate the assembly of the flagellum. The fT3SS and its interactions with various subunits also controls when subunits are exported for assembly. Before discussing how the fT3SS imposes order

on subunit export, it is important to understand the structure and proposed assembly mechanisms of the *Salmonella* flagellum.

1.7 Flagellar structure in Salmonella

The basic architecture of the flagellum is shared across bacterial species with adaptions made for Gram-positive and Gram-negative organisms and the periplasmic flagella of the spirochetes. The flagellum is formed of three helical substructures – the basal body, the hook, and the filament – requiring the export and assembly of over different 20 protein subunits. Each substructure has a unique capping protein that direct subunits to the site of assembly (125). The hook and basal body are often grouped as one larger substructure called the HBB (hook-basal body) as they assemble in the same phase.

Each substructure is responsible for a different function that, in a perfect example of the whole being greater than the sum of its parts, allows the bacterium to move through the environment. The *Salmonella* flagellum is presented as a model structure (Figure 1.4). Running through the three substructures is a narrow channel through which subunits are exported for their assembly at the distal tip of the nascent flagellum (126). Recent work using cryoelectron microscopy (cyroEM) revealed the rod and hook central channel, with a diameter of only 13 Å, to be narrower than that of the filament, with a diameter of 20 Å (Figure 1.5) (127, 128).

1.7.1 The basal body

The basal body spans the cell envelope, hosting the “drive-shaft” rod, a series of rings, and the motor-stator complex (126). The cytoplasmic (C) ring is a bell-shaped structure inside the cell made up of FliG/FliM/FliN, which, as previously discussed, serves as the motor and directional “switch” for rotation. The C ring is attached to the membrane-supramembrane (MS) ring, made up of the FliF subunit (126). The MS ring rests in a closed conformation until FliE is assembled, at which time the channel is opened to allow for proximal rod assembly (129). At the base of each flagellum, aligned with the export channel, is the dedicated flagellar type III secretion system (fT3SS), which sits inside the MS ring and bulges into the cytoplasm as a sort of cage. Type III secretion and its role in subunit assembly will be discussed in greater detail in

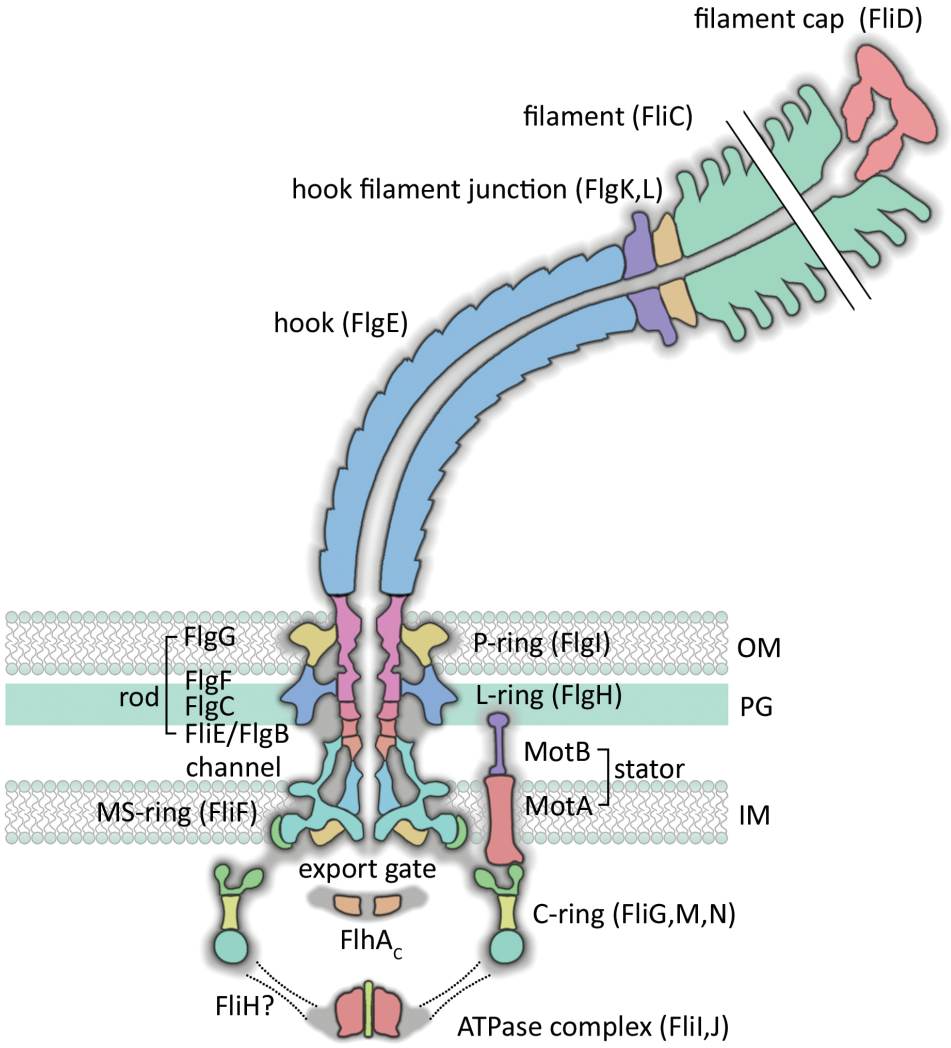


Figure 1.4 The structure of the flagellum is an elegant example of millions of years of evolution

Structure of the *Salmonella* flagellum, a model for Gram-negative bacteria flagella. The flagellum spans the cell envelope and extends far into environment, reaching lengths of 15-20 μm. Each subunit must be exported from inside the cell to the distal site of assembly through the narrow central channel. Figure adapted from Evans, Hughes, and Fraser, 2014.

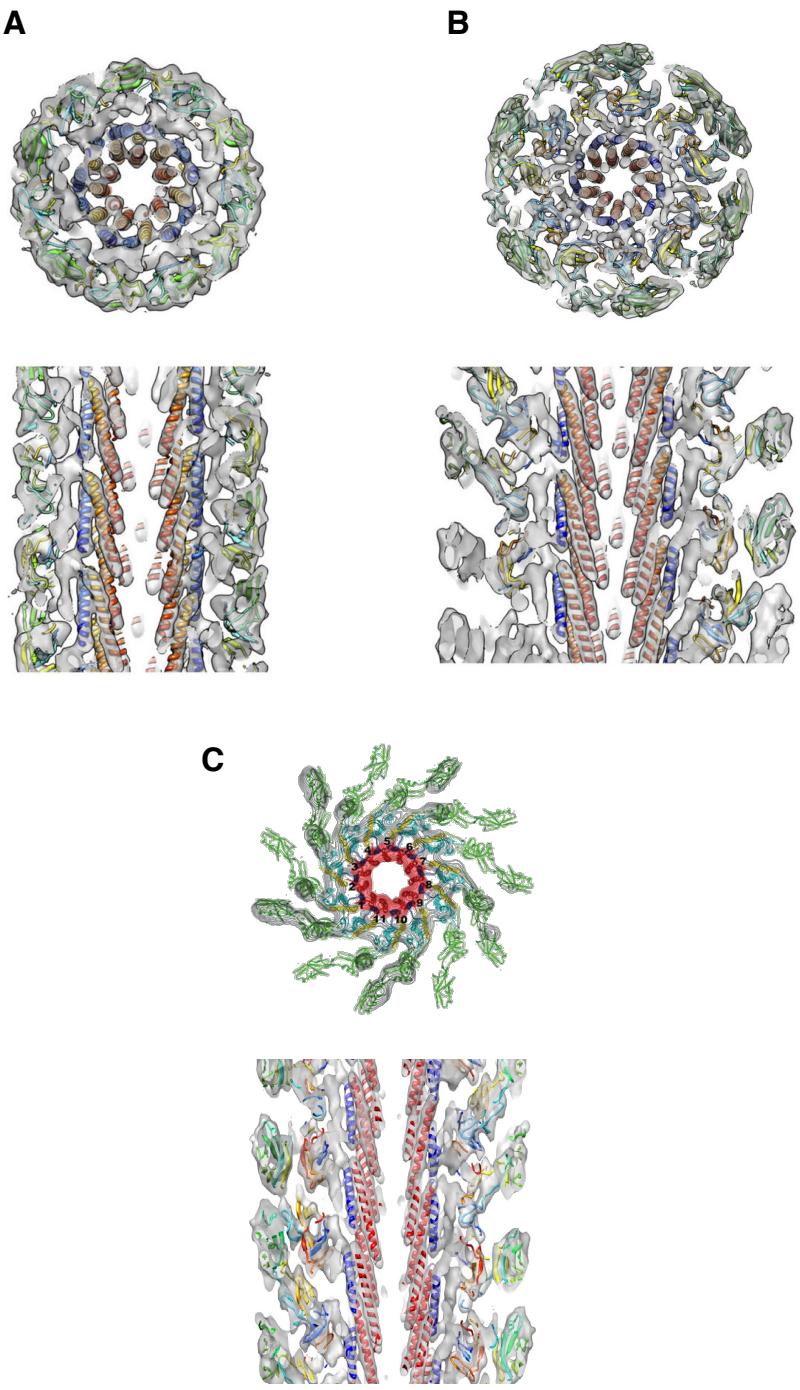


Figure 1.5 Flagellar subunits must travel through a narrow channel to reach the distal site of assembly

(A) The rod channel has a diameter of only 13 Å. Figure adapted from Fujii et al., 2017.

(B) The hook channel has a diameter of only 13 Å. Figure adapted from Fujii et al., 2017.

(C) The filament channel is wider than the rod and hook channel and has a diameter of 20 Å. Figure adapted from Yonekura et al., 2003.

the following sections, but, briefly, the fT3SS serves to unfold and export subunits through the narrow central channel of the flagellum and to separate subunits into two export phases, “early” and “late” (126).

There are two additional rings in Gram-negative organisms. The lipopolysaccharide (L) ring is made up of FlgH subunits and the peptidoglycan (P) ring is made up of FlgI subunits. These rings serve as bushings to allow the rod to rotate freely in the cell envelope (126). The gap between the distal rod and the L ring is not great, only around 10 Å, and it is proposed that the rod rotates freely without lubricant against the L ring due to repulsive forces between negative residues on the distal rod subunit, FlgG, and FlgH (128). As no atomic resolution structures for the P and L ring exist, this hypothesis is currently difficult to test.

The rod can be separated into two parts – the proximal rod, made up of 5-6 copies each of FliE, FlgB, FlgC, and FlgF, and the distal rod made up of between 26 and 42 copies of FlgG (128, 130). Unlike the other flagellar substructures, the rod is quite rigid because of the tight packing between the domains of FlgG during distal rod assembly (Figure 1.5A). This packing prevents flexing of FlgG, allowing the drive shift to remain rigid while rotating (128).

Rod assembly order is presumed to proceed FlgB, FlgC, and FlgF to first assemble the proximal rod. Only after the assembly of the proximal rod does the rod form the tube-like structure that allows for FlgG to polymerise into the distal rod (129). Thus, there is clear separation between proximal rod subunit assembly and that of the distal rod. But the precise order of these subunits appears to vary in different species.

Recent work has attempted to describe assembly order using a variety of methods. Zhao and colleagues used cryoelectron tomography to unpick the rod order in the spirochete *Borrelia burgdorferi*. Subsequent deletions of rod subunits *in vivo* revealed an order of FliE, FlgB, FlgC, FlhO (the spirochete FlgF homologue), and FlgG (129). *In vitro* work using rod subunits from *Salmonella* supports this order of rod subunits (131, 132).

Work in the Gram-negative bacterium *Rhodobacter sphaeroides*, however, found previously unknown *in vitro* interactions between FlgF and FlgB and between FlgC and FlgG. No direct interactions between FlgC and FlgF were detected. The authors propose FlgC as a checkpoint between the proximal and distal rod that assembles only upon FlgF polymerisation into the structure (133). Even though the exact export and assembly order of FlgC and FlgF may vary by species, these data support the need for cooperation between FlgB, FlgC, and FlgF (FlhO) to assemble the proximal rod (129, 133, 134).

What controls the transition between rod and hook assembly? A combination of physical length control and loss of the rod cap in Gram-negative organisms. FlgG polymerises until it reaches the outer membrane, the force of its polymerisation not enough to punch through the outer membrane (135). Next, the P ring assembles around the distal rod (128, 136). The L ring then assembles in the outer membrane, removes the rod capping complex, and allows the rod to span the cell envelope (136, 137). FlgD, the hook cap subunit, finally assembles on the completed distal rod to begin hook subunit polymerisation (137).

1.7.2 The hook

The second substructure, the curved hook, serves as the universal joint, transferring the torque generated by the motors to the helical filament (3, 125). The hook is made up of approximately 120 subunits, named FlgE (130). FlgE contains four domains – D0, Dc, D1, and D2 (138). Each domain contributes to the overall flexibility of the hook. The tightly packed D0 forms the inner core, lining the flagellar central channel, and provides mechanical stability. The coiled coil α-helices of the hook D0 domain are tilted to allow for extension and compression of the universal joint (127, 138). The Dc and D2 domains provide twisting rigidity and the D1 domain allows the hook to adopt highly curved conformations (Figure 1.5B) (139).

Whereas the rod and filament remain rigid during rotation, hook protofilaments roll during flagella rotation, occupying different places along the curve and changing their conformation and the distance between the repeating subunits (140). The 11 protofilaments shorten while inside the hook bend and extend once on the outside

(140). This innate flexibility is not vital to the swimming speed of the organism, but is critical to allow for flagella bundle formation in multi-flagellated bacteria (139).

Hook length in *Salmonella* Typhimurium is controlled at 55 nm ± 6 nm for efficient operation (141). This length is measured intermittently throughout hook assembly by a secreted molecular ruler protein called FliK (142). The length of the hook is tied to the export and assembly phases of the flagellum. Once a mature length is measured, early subunit export ceases and hook-junction proteins are exported to assemble on top of the hook to prepare for filament export and assembly. This results in an export specificity switch at the fT3SS that sees “early” subunit export give way to “late” subunit export (105). The export specificity switch will be discussed in greater detail in section 1.9.

Once the HBB is assembled, the “early” structure of the flagellum is considered complete and the “late” structure, primarily the filament, can begin assembly. Before the filament is assembled, however, the three hook-filament junction subunits must be exported and assembled. Two of these subunits, FlgK and FlgL, polymerise into a thin layer of approximately 11 subunits each, accounting for two turns of the helical structure (105, 143, 144). The third junction subunit is FliD, the filament cap subunit. FlgK and FlgL are the only assembled subunits that do not require a specialised cap to control their assembly into the distal hook. The reason why a cap is missing remains a mystery (125).

1.7.3 The filament

The filament represents the largest substructure of the flagellum, comprising tens of thousands of flagellin subunits, which account for ~1% of total cell protein. These subunits polymerise into supercoiled protofilaments that can extend up to 20 μm outside the bacterial cell (125, 126). Filament length does vary across the cell as longer filaments are often broken by physical forces, such as mechanical shearing, before being rebuilt by the cell (125, 145). The FliD cap is continuously exported, thus allowing for the rapid regrowth of sheared filaments (146). FliD assembles as a homo-pentamer at the distal end of the filament, forming a flat cap with five legs (147, 148). The leg domains can accommodate the different positions of individual flagellin

subunits as the helix grows outward, like stepping up a helical stairway. The FliD cap fills the place on the growing filament meant for next flagellin monomer. Once the next flagellin monomer reaches the site of assembly, the monomer replaces FliD and the legs of the cap move to the neighboring position where the process can restart once another flagellin monomer is delivered (125, 149). Each flagellin monomer is made up of four domains – D0, D1, D2, and D3. The D0 and D1 domains form the inner and outer tube, respectively, that creates the central channel through which subunits are exported. The D2 and D3 domains form the external surface of the filament (Figure 1.5C) (150–152).

As the motor rotates the entire structure, the filament moves through the liquid environment like an Archimedean screw to propel the cell (153). When multi-flagellated bacteria switch from a run to a tumble, the handedness of the 11 protofilament supercoils can switch between two straight filament types, left and right (L- and R-type), releasing the bundle quickly and stopping motion (154, 155). The repeat distances between L- and R-type protofilaments are similar, 52.7 Å and 51.9 Å, respectively. Therefore, this 0.8 Å difference requires flagellin to switch between types with a sub-ångstrom scale of precision (156, 157). Quite a remarkable feat!

This lengthwise mechanical switch is controlled by the outer-tube D1 domain. D1 domains interact within each protofilament, but not laterally across protofilaments. This allows each protofilament to switch between the L- and R-types independently, although cooperation is still needed among the protofilaments for a properly functioning filament (157). This cooperation comes from shearing forces between protofilaments caused by the twisting force generated by a switching motor complex. These shearing forces cause a conformational change at the D1 domain, causing them to slide upwards and to tighten the interaction between two of the domain's α-helices. These changes ripple through the D1 domains in the 11 protofilaments and result in the shift from L-type to R-type filaments (158). In such a manner, the otherwise rigid filament can change conformations to accommodate the different physical forces at work when the cell is swimming and when it needs to slow down or switch direction.

The structure of the flagellum is a beautiful example of a complex structure formed by inter- and intra-subunit interactions working in concert to self-assemble into a functioning nanomachine. Other than the initial structures that require the Sec pathway to integrate into the inner membrane, the structure is completely self-assembling through the fT3SS (125). Flagellin monomers, for example, must transit an extremely narrow channel that may be up to 20 μm long. How, then, does the cell power subunit transit? What steps are necessary to move a subunit from inside the cell to the distal site of assembly?

1.8 Mechanisms of flagellar assembly

Before subunits can transit through the channel, they must be prepared for export. This process can be separated into early subunit export of the rod and hook subunits and late subunit export, that of the hook-filament junction proteins and flagellin. Export of the early and late subunits is dependent on the fT3SS, which itself must be assembled within the MS ring.

The MS ring, the L and P rings, and the fT3SS are assembled via the Sec pathway, while the C ring self-assembles on the MS ring via interactions with FliG. These structures probably coordinate assembly amongst themselves to ensure proper spacing and base structure (125, 130, 159). The fT3SS contains six transmembrane proteins – FlhA, FlhB, FliO, FliP, FliQ, and FliR – and a dedicated ATPase made up of three subunits – FliH, FliI, and FliJ (126). The FliH/FliI/FliJ ATPase is homologous to the F_1 domain of the F_0F_1 ATP synthase (160). Once these structures are assembled, the real show can begin.

Flagellar subunits must be targeted in a step-wise process to the fT3SS, either through an N-terminal signal sequence or a specialised chaperone, in the case of late subunits (Figure 1.6) (84, 85, 161–165). First, the ATPase recruits the (chaperone-)subunit complexes to the appropriate export gate. The cytoplasmic domain of FlhA serves as the export gate for late subunits (126). FlhA forms a homononameric ring aligned with the central export channel (166). The central stalk of the ATPase, FliJ, interacts with the cytoplasmic domain of FlhA, providing a link between the export gate and the ATPase, and acts to bring chaperones to the export gate (76, 167). There at the gate,

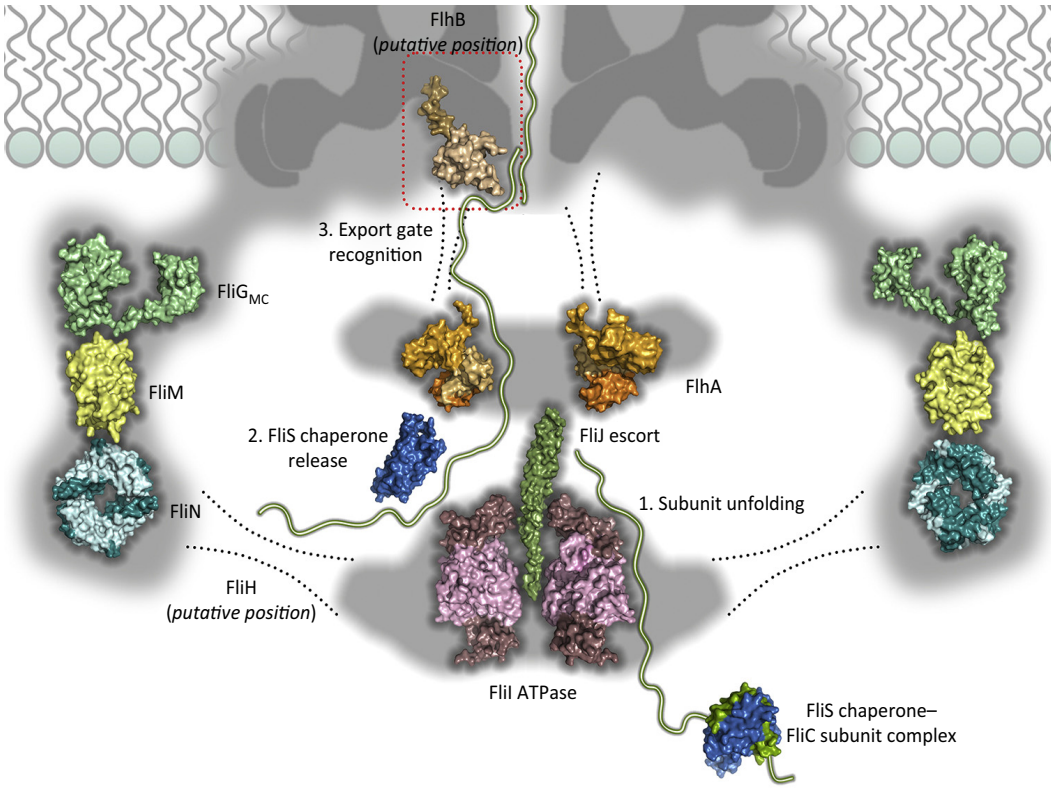


Figure 1.6 Flagellar subunits interact with the fT3SS in a step-wise process prior to export

The flagellar type III secretion system (fT3SS) shown in an active state with the ATPase (FliHIJ) engaged with one of the export gate components, FlhA. Flagellin (FliC) provides an example of the steps chaperoned subunits must go through prior to export. Figure adapted from Evans, Hughes, and Fraser, 2014.

the ATPase can facilitate the release of chaperone and the unfolding of subunit prior to export (126). In the case of late subunits, affinity of chaperone-subunit complex for their export gate reflect the order of the final structure. This gate-chaperone-subunit interaction may serve as a checkpoint to impose order onto the late structure, preventing export of flagellin before the hook-filament junction proteins have been exported (168).

The export gate FlhA complex is a proton-protein antiporter that can make use of the PMF to translocate protein across the cell envelope. FlhA contains a charged residue, Aspartate-208, on its transmembrane α-helix$_5$ that is directly involved in protein export (169). The efficiency is increased in the presence of the ATPase, suggesting the need for two energy sources (170).

Early subunits are not chaperoned to the export gate, but rather possess a secretion signal that allows them to bind directly to their export gate, FlhB (165). This signal and the role of FlhB will be discussed in section 1.9.1. Early subunits are still exported at least partially unfolded, although FlhB does not seem to have a proton channel in its transmembrane α-helical domains as FlhA does. Most research into early subunit export has studied how early subunits contact FlhB. Early subunit interactions with the ATPase, if any, remain unknown.

A picture emerges of subunits reaching the fT3SS, interacting first with the ATPase before being handed off to the export gate. How are subunits exported through the channel? For some time, it has been shown that subunits must be at least partially unfolded by the PMF with help from the ATPase before being exported (75, 90, 91). The central channel diameter, 13 Å for the rod and hook and 20 Å for the filament, is not wide enough to accommodate folded protein, although it is predicted that limited secondary structure may be able to pass through (127, 128). Radics and colleagues used cryo-EM to investigate the structure of subunits during export through the *Salmonella* injectisome T3SS. Their work mapped the path an exporting subunit follows and found that subunits must be unfolded during export to follow the export path (68). As the flagellum has a narrower channel than the injectisome, flagellar subunits are probably also exported fully unfolded.

Now that the subunit has reached the fT3SS and docked at the gate, how does it reach the distal site of assembly? Especially as there is no apparent source of energy in the flagellar channel to power subunit transit (171). Recent work suggests two mechanisms.

Evans and colleagues proposed that subunits travel to the site of assembly as a subunit chain (Figure 1.7A). Subunits link into chains via their termini as parallel α-helical coiled coils. The N-terminus of an export gate docked-subunit can be captured by the C-terminus of a channel subunit. Subunits then transit through the channel without the addition of an external energy source because of the inherent entropic force of the unfolded subunits. Once they reach the site of assembly, subunits polymerise into the growing structure, anchoring the chain. This anchor provides directionality to the chain, pulling it from the fT3SS at the base to the distal tip. As each subunit assembles into the growing structure, the subunit chain tightens. The chain can then generate enough force to pull the next docked subunit off the export gate, which causes the chain to relax and restarts the process. In this way, the chain tightens and relaxes in rapid succession as subunits polymerise into the structure, capable of delivering subunits to the site of assembly at a constant rate (165). The chain mechanism, then, explains another recent observation that flagellum growth rate is independent of length, something diffusion alone cannot account for (165, 172).

For the chain mechanism to work, there are a few conditions that must be met. Subunits must be able to link head-to-tail via their termini, this complex must be able to move through the channel, and subsequent forces forming the chain must be strong enough to hold the chain together but weak enough to break upon subunit polymerisation into the structure.

The authors showed by *in vitro* capture assays and *in vivo* crosslinking experiments that early and late subunits can link up by their termini, and that this linkage was specific to the termini. Importantly, they showed that different early subunits could link their termini, leading to the possibility of a mixed subunit chain. Late subunits could not capture early subunits into a chain, so a mixed export “phase” chain is not possible.

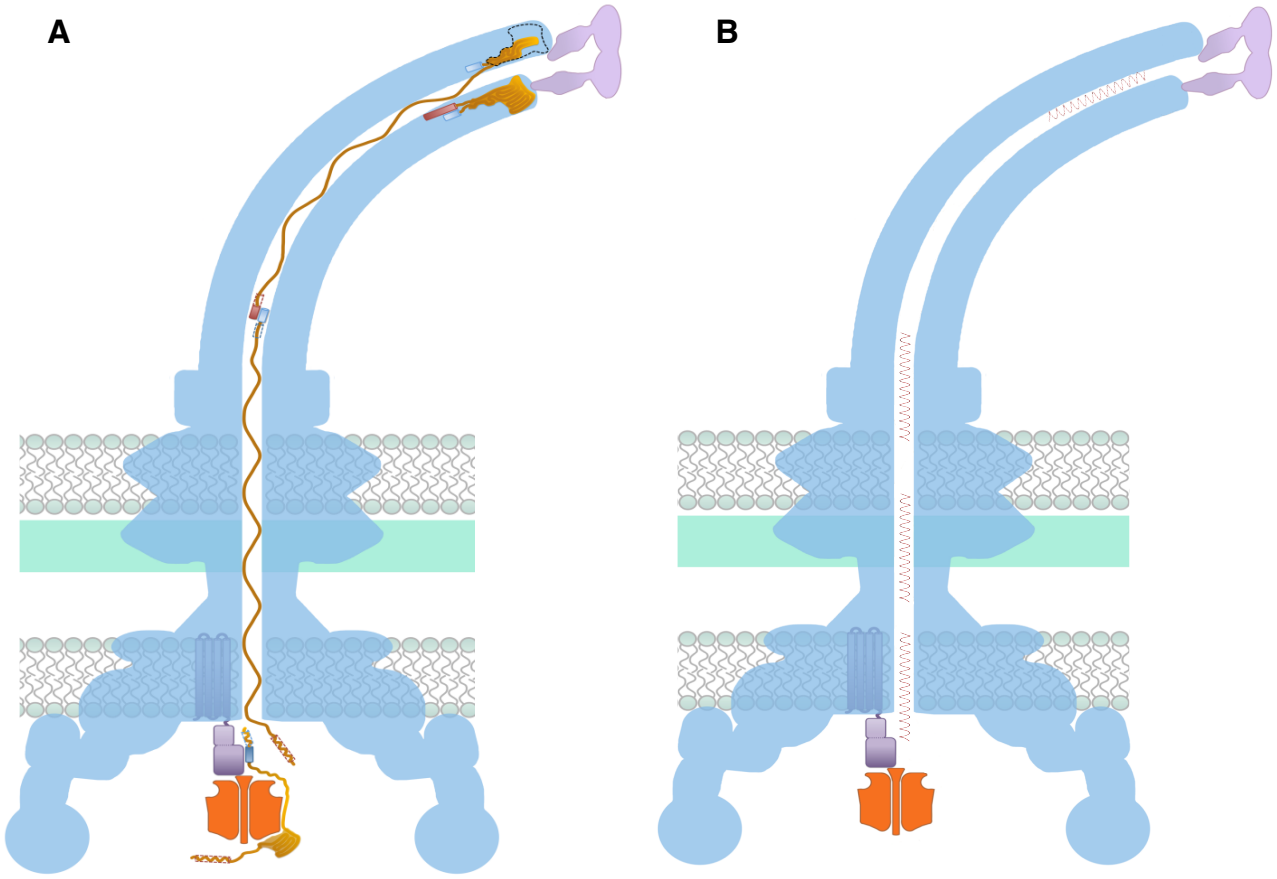


Figure 1.7 Mechanisms for subunit transit and flagellum growth

(A) The proposed chain mechanism for flagellum assembly. Subunits are unfolded at the fT3SS, leaving the terminus of a docked subunit free to interact with a channel bound subunit to form a coiled coil domain. These head-to-tail links form a subunit chain that travels the length of the structure. Transit is self-powered, the entropic force of the subunits enough to power transit, and unidirectional as subunit folding into the structure or the environment prevents the chain from moving back into the cell. Upon a chain break, the next subunit can diffuse to the site of assembly to restart subunit chaining. The subunit chain is capable of delivering subunits at a constant rate, independent of flagellum length (Evans et al., 2013). Figure adapted from Evans et al., 2013.

(B) The injection-diffusion model of subunit transit. Subunits are injected into the channel by the fT3SS using energy provided by the PMF. Subunits then travel through the channel via single-file diffusion. The channel is too narrow to allow for overlapping interaction, so any subunit-subunit interactions are primarily ballistic. Growth slows as the length of the filament grows (Renault et al., 2017). Figure adapted from Evans et al., 2013.

They also showed that subunits lacking termini could not pull docked subunit from the export gate (165).

The subunit chain relies on three core forces – the force holding the subunit to the export gate (F_M), the force between the subunit termini forming the links of the chain (F_L), and the anchoring force as subunit polymerises into the flagellar tip (F_A). FlgD was found to bind the FlhB export gate weakly, with a K_d about 40 μM, resulting in a weak F_M of about ~30 pN for FlgD. This force is weaker than F_L, which is predicted to be between 400-700 pN. Finally, the F_A is at least an order of magnitude greater than F_L, approximately 11 nN when subunit polymerises into the structure. The F_A is weaker when subunit refolds in the external environment, only ~6 nN, but it is still strong enough to pull the chain through the channel. Each of these forces are successively stronger than the last, meaning each successive force can break the last but cannot be broken by the prior force, allowing the chain to move quickly and unidirectionally through the channel (165).

Unfortunately, Evans and colleagues did not discuss the width of the proposed coiled coils and whether they could fit through the channel. Another parallel coiled coil dimer found in nature is the leucine zipper. These zippers have a width of ~30 Å, much too large to fit through the HBB or filament channel. Leucine zippers are twisted, which causes them to bulge (173). It is reasonable to assume that straightened parallel coiled coils as proposed by Evans and colleagues would be narrower. Another example of the structure subunit termini may adapt is an α-helical hairpin. These hairpin structures are still wider than the narrowest part of the channel, just over 14 Å at their narrowest point (174).

How could termini interact in the channel? That question now becomes more important to the viability of the chain mechanism. The two examples provided aren't of flagellar subunits, so it is possible that flagellar subunit termini make different coiled coil structures. Even if narrower than the channel, there is not much room for overlapping domains to pass through. The wider the coiled coils are, the more likely they would get stuck in the channel, requiring a very large force to push or pull them through, or could interact with the residues lining the channel and possibly slow down the chain.

Of course, this assumes termini do adopt secondary structure in the channel, which may not be the case. The termini structure of flagellar subunits during export needs to be investigated further to determine if they can fit through the channel.

If not a chain mechanism, then what? Renault and colleagues present evidence for an injection-diffusion model of subunit transit (Figure 1.7B). In their model, subunits do not form a chain to transit outside the cell. Rather, they propose subunits are injected into the channel with energy provided by the PMF and diffuse individually to the site of assembly. They note that the growth rate of filament length is not independent of the length of the filament by elegantly labelling *in situ* different stages of filament growth in living cells. During flagellum growth, they exchange the fluorescent dyes to detect the different filament fragments. This allowed Renault and colleagues to track filament growth accurately and precisely over time. They watched the elongation rate hasten when filaments were made shorter by mechanical forces. They conclude that the growth rate is inversely proportional to filament length and slows drastically from an initial rate of roughly ~83-100 nm min^{-1} until the growth rate becomes negligible in very long filaments. When they disrupted the PMF by using carbonyl cyanide *m*-chlorophenyl hydrazone (CCCP), they found the growth rate slowed to ~18 nm min^{-1} and became linear, similar to the previously report independent growth rate of ~13 nm min^{-1} (99, 172). Thus, they concluded that PMF is used to inject unfolded monomers into the channel (99).

How may subunits diffuse to the site of assembly? In the injection-diffusion model, subunit monomers must diffuse in single file. The authors reason the channel is too narrow to allow for end-to-end binding, so they predict any interaction between monomers to be ballistic. As the channel fills with monomers, diffusion would slow accordingly because monomers would be unable to pass one another in the channel (99).

These two models do leave questions unanswered. Can subunits chain? It certainly seems so for early subunits, where robust *in vivo* and *in vitro* evidence was presented. For late subunits the picture is murkier (165). There is other evidence that a subunit chain may not drive subunit transit. Hook-filament junction proteins (FlgK, FlgL) and

the filament cap (FliD) are secreted in relatively high levels throughout filament assembly (175). But flagellin does not appear to bind to the hook-filament junction proteins. It can, however, bind to FliD, which can bind to the junction proteins (176). It is possible, then, that if the subunit chain is broken by a junction subunit, FliD can restart the chain by acting as a bridge between flagellin and the junction proteins. The chances of this would depend on subunit expression level and the export gate's ability to discriminate subunits.

Can subunits interact in and with the channel? The injection-diffusion model makes a fair assumption that the channel is too narrow to allow for any monomers to interact end-to-end during transit (99). No one has observed a subunit trapped in the flagellar channel, so it is hard to conclusively say what structure subunits adopt in the channel. But, if subunits were to interact in the channel, they could act as a cork and prevent subunit travel, effectively killing that flagellum until the interaction weakened or the cell spent enough energy to push the subunits through. In a long flagellum, this would grow costly. The chain mechanism pre-empts this by claiming that subunits must interact to transit to the site of assembly. Although both models support that friction is not the rate-limiting step in pulling subunits, they do not discuss in detail whether subunits might interact with the channel wall at any step (99, 165). If subunits travel through the chain unfolded, as is the case in the injectisome, that would leave all their residues to interact with available channel residues (68). In this case, subunit transit could be significantly slowed depending on the strength and kinetics of these interactions.

As with so many things in nature, the true subunit transit mechanism is probably closer to a combination of both the chain mechanism and the injection-diffusion mechanism. In fact, Evans and colleagues proposed diffusion as the necessary restart whenever the chain breaks due to inability of subunits to chain or a sheared filament necessitating a new cap (165). This could also explain why, although the chain is capable of delivering subunits at a constant rate, this would not be the case in nature (99, 165). So, although these two models are much more advanced than past proposals, more research is needed into subunit structure in the channel and inter-subunit interactions to elucidate the true mechanism of subunit transit.

Subunit transit and assembly are only part of the story, of course. Flagellar structure is tightly ordered by a combination of gene expression hierarchy discussed in section 1.6 and protein-protein interactions at the fT3SS. The fT3SS imposes a strict order on subunit export and assembly via many interactions between fT3SS components and other flagellar subunits. This control is tied to gene expression via the fT3SS monitoring the phases of flagellar assembly. Without this level of feedback and control between the export machinery and gene expression, a competent structure cannot be produced. The following section discusses the roles of a few fT3SS components in imposing order on subunit export and assembly.

1.9 Ordered assembly

A shared and notable feature among T3SSs is the ability to discriminate between subunits and impose order onto subunit export. T3SS systems operate in phases – an early phase and a late phase for the fT3SS and an early, intermediate, and late phase for the injectisome. The early phase sees the export and assembly of basal body and hook subunits in the flagellar system and basal body and needle structure in the injectisome system (105, 177). The length of this early structure is measured by a secreted molecular ruler (178–181). Once a mature hook or needle length has been measured, a signal is sent to the export machinery, leading to a switch in export specificity to the second hierarchy of subunits (105, 177). Late flagellar subunits, those that assemble into the filament, or intermediate injectisome subunits, those that make up the translocon, are now made ready for export (126, 182).

The export specificity switch remains one of the greatest outstanding questions in type III secretion. Much work over the past few decades has uncovered two key proteins for the switch found in T3SSs – a membrane-anchored subunit export gate, FlhB/SctU, and the aforementioned secreted molecular ruler, FliK/SctP. How do these proteins work together to cause the switch?

A brief note on nomenclature. Unlike flagellar subunits nomenclature, injectisome subunit nomenclature is not unified. There are efforts by some groups to unify injectisome nomenclature using the *s*ecretion and *c*ellular *t*ranslocation (Sct) nomenclature (183). In the following sections, injectisome components will be

identified with their Sct names unless otherwise specified. Where both names are used (e.g. FlhB/SctU), the information pertains to both T3SSs.

1.9.1 The early subunit export gate, FlhB

The early subunit export gate FlhB/SctU is a core member of the dedicated type III secretion export apparatus (Figure 1.8A-D), integrated into the membrane via the Sec pathway. FlhB/SctU contains two domains – a transmembrane domain comprising four alpha-helices through the inner member and a cytosolic domain (126). These two domains are connected by a linker (184). Although usually considered disordered, the linker domain of *Yersinia* SctU was shown to adapt a helical conformation when near a negatively-charged membrane, directly interacting with the membrane via a lysine and arginine on the membrane-face of the helix. Mutating the positively-charged, solvent exposed residues in the helix did not affect the interaction with the membrane, but affected effector molecule secretion, suggesting that this helical linker may serve to position the early subunit export gate to function appropriately (185).

One copy of *Salmonella* pathogenicity island 1 (SPI-1) SctU was found at the export machinery by using mass spectrometry of peptide concatenated standards and synthetic stable isotope-labeled peptides (186). Genetic fusions observed in the flagellum of other species (like the *Clostridia*) and biochemical work completed in *Salmonella* show that a fusion between *flhB* and another export apparatus protein, dubbed *fliR* in flagella, is motile. These data support the hypothesis that FlhB is located in association with FliR near the export channel at the inner membrane (186, 187). Purification of full length FlhB and HBB showed that FlhB is able to dimerise, certainly *in vitro* in detergent micelles and possibly *in vivo* (188). The FlhB cytoplasmic domains were not found to interact, suggesting that the transmembrane domains were responsible for dimerization, if it were to occur at all in the cell, and may force the cytoplasmic domains into close proximity (188, 189). Further, deletion of *flhB* did not substantially alter the T3SS as measured by electron cryo-tomography (166). These data suggest that only one copy, two at most, of FlhB/SctU are present within the T3SS, although it may be system-dependent.

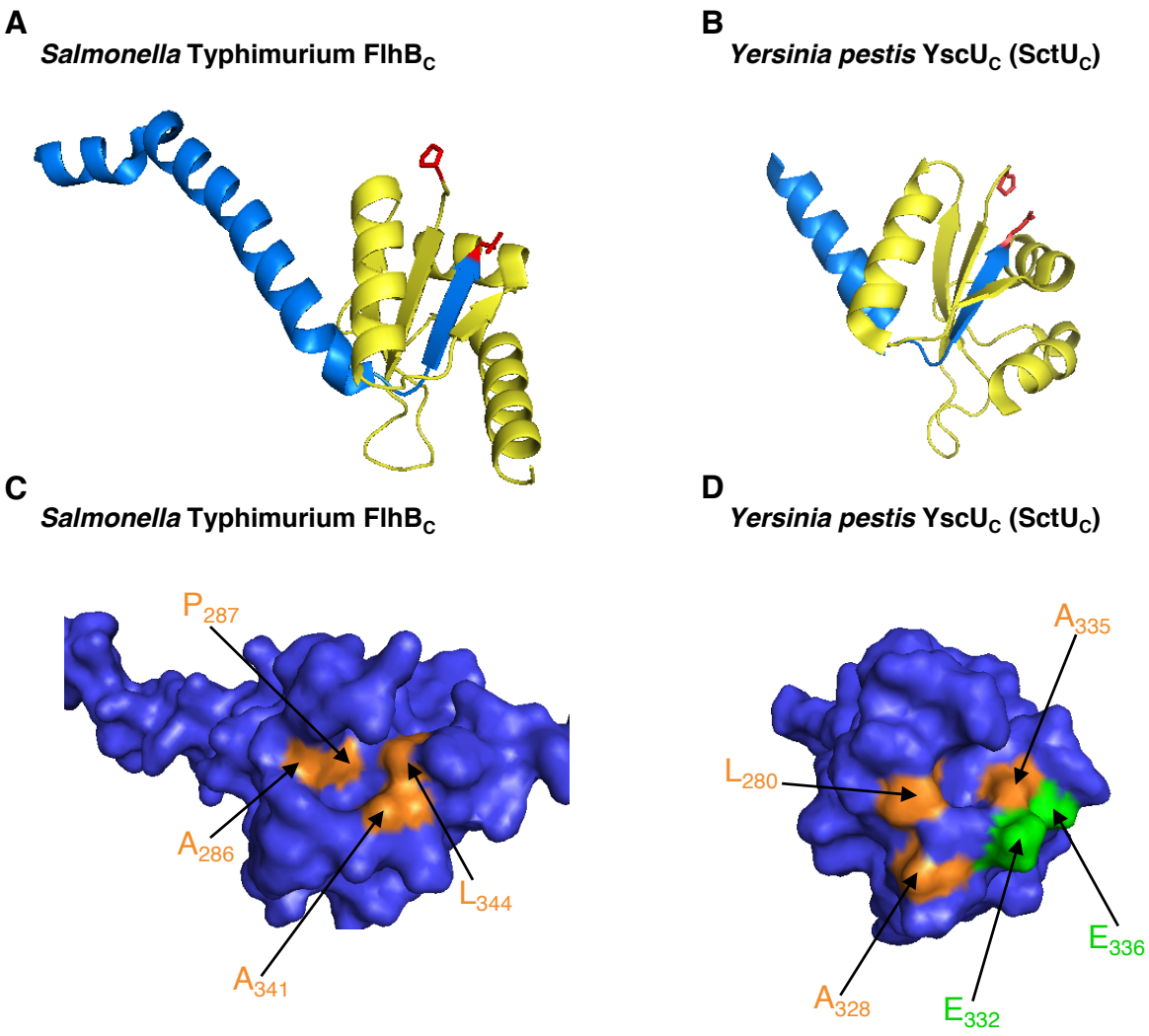


Figure 1.8 The early subunit export gate and secreted molecular ruler share features in both flagellar and needle T3SSs

(A) Atomic resolution structure of autocleaved *Salmonella* $FlhB_C$. $FlhB_{CN}$ is in blue, $FlhB_{CC}$ is in yellow. The autocleavage site is in red.

(B) Atomic resolution structure of autocleaved *Yersinia pestis* $YscU_C$ ($SctU_C$). $YscU_{CN}$ is in blue, $YscU_{CC}$ is in yellow. The autocleavage site is in red.

(C) Surface-exposed hydrophobic patch where flagellar rod and hook subunits are proposed to bind prior to export (Evans et al., 2013).

(D) Surface-exposed binding site for the *Yersinia* molecular ruler YscP (SctP) (Ho et al., 2017).

(E) Sequences of putative GRMs of type III molecular rulers that may bind to FlhB/SctU. Sequences in red are confirmed to interact with the early subunit export gate, FlhB/SctU.

The cytosolic domain of FlhB/SctU ($FlhB_C$/$SctU_C$) contains a highly conserved NPTH loop found in both the flagellar and injectisome T3S export apparatus. This loop undergoes rapid succinimide-mediated autocleavage with a half-life of approximately five minutes, resulting in two, non-covalently interacting cytoplasmic domains – the cytosolic N-terminal domain ($FlhB_{CN}$/$SctU_{CN}$) and the cytosolic C-terminal domain ($FlhB_{CC}$/$SctU_{CC}$) (Figure 1.8A and 1.8B) (184, 189–191). After autocleavage, the two polypeptides remain tightly associated with one another, although the exact magnitude of this association is unknown (190, 192).

The autocleavage of the early subunit export gate is necessary for the proper functioning of the type III export machinery. A single residue mutation at the FlhB NPTH loop resulting in a loss of or inefficient autocleavage leads to the polyhook phenotype, where the cell continues to export and assemble flagellar hook subunit instead of controlling hook length. These flagella are therefore locked into early subunit export and are not motile (193). Similar results have been shown for SctU in a variety of pathogens (194–197). Autocleavage-deficient mutants of *flhB* are, however, still capable of exporting early subunits (193). Autocleavage appears only critical, then, to switch export specificity to intermediate injectisome or late flagellar subunits.

As previously discussed, early flagellar subunits are not chaperoned to the export machinery. Rather, they possess an N-terminal, hydrophobic, five-residue secretion signal, dubbed the gate recognition motif (GRM) (86). This motif has also been reported in the injectisome molecular ruler, an early subunit secreted during needle assembly in the injectisome T3SS (Figure 1.8E) (88, 89). $FlhB_{CC}$ contains the putative surface-exposed, hydrophobic binding pocket for the early subunits (Figure 1.8C) (86). Recently, a similar binding pocket was uncovered for needle rod subunits in *Yersinia pseudotuberculosis* $SctU_{CC}$ (Figure 1.8D) (89).

What might happen to FlhB after the switch? In *Yersinia*, it has been shown that $SctU_{CC}$ can dissociate from the rest of the SctU export gate after autocleavage and that it may be secreted from the cell after the switch is triggered (198). It was later determined that this domain may carry a secretion signal in the last 15 residues (199).

Whether $SctU_{CC}$ is secreted into host cells has not been investigated, so its ultimate fate is unknown. It also remains unclear if other FlhB/SctU-type proteins share this feature.

1.9.2 The molecular ruler, FliK

In addition to the early subunit export gate, the secreted molecular ruler, FliK/SctP, is required to switch export specificity at the export machinery. The type III secreted molecular ruler is classed as an early subunit, possessing a GRM like other flagellar subunits, but does not assemble into the final structure (Figure 1.8E) (86, 88, 89). Instead, the N-terminal domain of FliK measures the length of the nascent hook or needle structure before being secreted into the environment (200).

The C-terminal domain of FliK/SctP was dubbed the type III secretion substrate specificity switch (T3S4) (201). Genetic studies confirmed that 50 residues located in the core T3S4 domain and the last five residues of FliK were critical to the export specificity switch. These deletions had some effect on expression level, but did not affect export of the molecular ruler (202).

FliK/SctP share low sequence homology across species, but do share structural and functional features. FliK/SctP proteins possess a “ball and chain” structure, consisting of an extended N-terminal ruler domain and a folded C-terminal T3S4 domain (88, 89, 203). These domains are connected by a disordered linker, which can be lengthened or shortened to correspondingly lengthen or shorten the flagellar hook (204). Thus, the length of the hook is not dependent on any inherent polymerisation feature of FlgE, like is the case with the distal rod, but under the direct control of the length of the molecular ruler’s N-terminus and linker. Atomic force microscopy revealed that the FliK N-terminal domain could form a stable ball in solution, although it is not structured enough for structural studies using NMR or crystallography (205). This suggests that the N-terminal domain of the molecular ruler is easily unfolded by the type III export machinery prior to export so that the unfolded N-terminus can serve as a measuring stick as it transits through the export channel.

The number of molecular rulers secreted by the injectisome or the flagellar T3SS varies. In the injectisome, SctP is secreted once; thus, a single N-terminal domain of the molecular ruler must measure the growing needle and trigger the export specificity switch at the injectisome export machinery (206). *Yersinia* SctP was observed to have two secretion signals in the N-terminal domain, supporting the notion that the N-terminus is secreted first (207). The *Yersinia* molecular ruler has been observed at the tip of the needle by immunogold electron microscopy (208), hinting that the single molecular ruler may continue to associate with the needle as it is assembled.

In contrast, FliK is secreted intermittently over the course of hook assembly (142). Unlike the injectisome which does not have a needle cap, FliK can interact *in vitro* with the hook cap, the protein found at the tip of the nascent hook (209). The N-terminal domain of the molecular ruler must follow the growing hook throughout assembly before being secreted into the environment by the hook cap, again supporting that the N-terminal domain is secreted first. To study the distribution of FliK prior to export, Hirano and colleagues fused the cyan fluorescent protein (CFP) to the N-terminus of the FliK molecular ruler. Although this fusion may be assumed to have masked the FliK N-terminal GRM, they found that cells expressing the CFP-FliK fusion swam comparable to wild type. Their results demonstrated that FliK is homogenously located throughout the cell after expression and that it presumably only interacts with the T3SS in preparation of export (210). Therefore, the secretion of the FliK N-terminal domain serves to place the FliK C-terminal T3S4 near the export machinery long enough to lead to the export specificity switch.

It is well-studied that the early subunit export gate and secreted molecular ruler must communicate to switch export specificity at the type III export machinery. This has been observed from genetic studies and biochemical studies using full length protein (88, 89, 191, 202, 204). Genetic studies suggest the C-terminal T3S4 domain of the molecular ruler may interact with the cytoplasmic domain of the early subunit export gate to cause the export specificity switch (202). But subsequent biochemical analysis of an interaction between $FliK_C$/$SctP_C$ and $FlhB_C$/$SctU_C$ are crucially lacking. Only a handful of studies have investigated the possibility of the molecular ruler T3S4 binding *in vitro* to the early subunit export gate and have found mixed results (86, 88, 89, 211).

Lorenz and colleague found that SctP N-terminus of *Xanthomonas* was unable to bind SctU without the C-terminal T3S4 present. They did not investigate if the T3S4 could bind the SctU on its own (211). Evans and colleagues found that FliK without the GRM ($FliK_{\Delta GRM}$) can bind GST-$FlhB_C$ weakly *in vitro* (86). Although, as $FliK_{\Delta GRM}$ is missing only five residues from the N-terminus but is otherwise intact, it is difficult to determine if this weaker binding was due specifically to the $FliK_C$ T3S4 domain associating with $FlhB_C$. Ho and colleagues found that *Yersinia* $SctP_C$ only associated with $SctU_C$ when the linker region was present with the C-terminal T3S4 region. The C-terminal T3S4 region alone did not show evidence of an interaction, suggesting it is either very weak or non-existent (89). Similarly, in *Pseudomonas*, the $SctP_C$ T3S4 alone was unable to bind $SctU_C$ *in vitro* (88).

Most biophysical studies measuring the affinity between the molecular ruler and early subunit export gate have used full length protein, rather than isolating individual domains for study (89, 191). It is therefore reasonable to suspect that these affinity measurements mostly relate to the GRM secretion motif near the molecular ruler's N-terminus. These data should not serve as demonstrative proof that the C-terminus binds to the early subunit export gate to trigger the export specificity switch, a hypothesis that remains in dispute. Further investigation should seek to isolate the domains of the molecular ruler to determine what domains may bind the early export gate and what the magnitude of binding may be.

1.9.3 Other proteins contributing to the switch

Whatever the interaction between FliK/SctP and FlhB/SctU may be, they are by far the most conserved players of the export specificity switch. But other proteins may play a limited role, depending on the species.

Some organisms, like *Salmonella*, possess a "cork" at the base of the flagellum, a protein called Flk (pronounced fluke) that may work in association with FlhA to prevent premature switching by FliK (212). It has been proposed that $FliK_C$ could only interact with $FlhB_C$ once it has bypassed the Flk cork, as would happen when FliK is exported through a mature hook (105). The location of Flk in the export machinery has not been determined, so it is hard to say if this may be the case. We do know that Flk faces the

cytoplasm, anchored via a C-terminal hydrophobic transmembrane tail, and prevents premature switching (213). Flk seems to manage the export specificity switch by coupling translation of *flgM* mRNA with P- and L-ring assembly (214). As Flk is found in only a handful of organisms, it remains poorly studied. But since it provides some control over the export specificity switch in *Salmonella* and other enteric bacteria, further investigation is merited.

FlhE is a conserved flagellar subunit in *Salmonella* and other peritrichous bacteria that is required for swarming motility, but not swimming motility (215). As it is not involved in flagellar biogenesis and found only in flagellar systems in a limited number of species, FlhE is also not a well-studied protein and it is the last member of the *flhABE* operon without a clearly established function. The loss of FlhE causes a proton influx into the cytoplasm, suggesting that it may serve to modulate proton flow through the fT3SS (216). When FlhE was deleted from a *Salmonella* strain defective in P- and L-ring assembly, the export specificity switch still occurred (212). At the moment, FlhE's exact role in the switch is pure conjecture. The structure of the protein has been solved, so further biochemical and biophysical work on FlhE is possible to elucidate its true function in flagella regulation (217).

Mutations in other proteins, such as FlhA, FlgG, and the C ring, have also been shown to affect the export specificity switch (212, 218). These mutations probably interfere with the FlhB-FliK control of the export specificity switch, either directly or by changing flagellar structure to allow for spontaneous switching.

1.9.4 State of the export specificity switch

The beginnings of a mechanism emerge when we piece together individual data on the early subunit export gate and secreted molecular ruler. The most basic mechanism requires the export and assembly of hook or needle subunits. During this stage, the molecular ruler binds to the autocleaved early subunit export gate through an N-terminal, hydrophobic gate recognition motif, the GRM. The molecular needle is secreted to measure the length of the nascent hook or needle. Once a mature length has been determined, the C-terminal T3S4 domain triggers the export specificity switch at the early subunit export gate. These facts are not disputed. But, due to some

quirks between the two type III systems and some gaps in our knowledge, the exact mechanism remains unknown. There are two recent hypotheses that seek to collate the most robust evidence for the export specificity switch in each type III system.

Erhardt and colleagues proposed a mechanism for the flagellar export specificity switch in 2011. In their model, the molecular ruler is secreted intermittently during hook assembly. In short hooks, the molecular ruler is secreted into the environment where it begins to refold, pulling the rest of the protein with it, before the T3S4 domain can trigger the switch at the early subunit export gate. This process continues until the hook has reached the desired physiological length. When another molecular ruler molecule is secreted to measure the completed hook, the interactions between the molecular ruler and assembled hook subunits slows secretion long enough to allow for the C-terminal T3S4 domain of the molecular ruler to trigger the switch at the early subunit gate (142).

In the injectisome, Bergeron and colleagues proposed a model that saw the T3S4 domain of *Pseudomonas* $SctP_C$ anchor at the tip of the nascent needle and the $SctP_N$ domain remain associated with the early subunit export gate. In this model, the disordered N-terminal domain is stretched while within the channel as the needle assembles. Once the ruler measures a mature needle length, the N-terminus domain docked at the early subunit gate triggers the switch, dislodges from the needle tip and the export gate, and is fully secreted (88).

This latter model has been challenged for two main reasons. The first, affinity measurements between the molecular ruler GRM and the early subunit export gate are micromolar, strongly suggesting transient interactions (89, 191). The affinity between *Yersinia* SctP and $SctU_C$, for example, was calculated by NMR to be 430 ± 293 μM *in vitro*. This interaction alone is doubtless not strong enough to hold the SctP N-terminus at the early subunit export gate as Bergeron and colleagues proposed, unless, *in vivo*, there are further partners yet to be discovered or the membrane is involved in stabilising the interaction (89).

Secondly, the flagellar and needle channel diameter is narrow, on the order of 20-25 Å at its widest (88, 127, 219). This channel only allows for unfolded subunits, and possibly partially unfolded subunits, to transit through. Observing subunits in the channel is technically challenging, but recent work has demonstrated the physical limitations of having multiple parallel subunits in the channel at any one time (68, 86, 99, 220). Even if SctP were not structured in the needle channel, it could interact with moving subunits and pose another barrier to efficient subunit transit.

It should be noted that the molecular ruler may not need to be secreted to trigger the export specificity switch. In one published example, hook length was indirectly controlled by the number of available C-terminal domains of FliK at the flagellar export apparatus, with overexpression of the truncated protein ultimately leading to near wild-type motility in the absence of any hook length measurement (210). This argues for a temporal element to the FliK-FlhB interaction that causes the export specificity switch. $FliK_C$ must be able to interact with $FlhB_C$, or trigger other interactions, just long enough to lead to the subunit switch. It's probable the timing of this interaction is equal to the rate of travel of $FliK_N$ through a mature hook. The FliK truncation, residues 100-405, that can switch export specificity in the absence of secretion left much of the FliK N-terminus intact. None of the truncations without the N-terminus were capable of secretion or switching (210). This truncation may, then, be sufficiently long enough to correctly time the interaction between the FliK T3S4 and $FlhB_C$ to trigger the switch. What further role, if any, the FliK N-terminus and disordered linker region plays in the export specificity switch besides measuring hook length is unknown.

It is reasonable to suggest that the two systems will have slightly different mechanisms because of the difference between secretion patterns of the molecular ruler, the additional phase of assembly in the injectisome, and the possibility that the fT3SS has two copies of the early subunit export gates instead of one as was found in the injectisome. Additionally, the export specificity switch often requires the orchestration of other cytoplasmic or membrane-anchored proteins beyond the molecular ruler and early subunit export gate, although that is highly system- and species-dependent. The shared evolutionary history between the flagellum and injectisome means that the two

mechanisms will have much overlap, even when accounting for species-specific idiosyncrasies.

1.10 Aims

Many bacteria, including pathogens like *Salmonella*, require type III secretion in order to colonise a host and cause infection. What is clear about the nanomachines utilising type III secretion is the amount of coordination extending from protein-protein interactions to controlled gene expression needed for the final product. This study will investigate some of the protein-protein interactions that must occur if a bacterium hopes to move through its environment to find a suitable site of infection or an environment better suited to growth and proliferation.

The fT3SS export specificity switch divides subunit export, and therefore overall assembly, into two phases. It intimately ties gene expression to flagellar substructure assembly through secretion of FlgM only after FliK has measured a mature hook and triggered the switch at FlhB. But it remains unclear how the export specificity switch occurs, particularly what conformational changes FliK may activate at FlhB to switch export specificity.

This study seeks to further examine FlhB's dual role as an early subunit export gate and the arbiter of the export specificity switch. *In vivo* techniques were used to investigate mutational effects on swimming motility, subunit export, and the fate of proteins in the cell in order to understand the interactions between early subunits and FlhB. *In vitro* techniques, such as circular dichroism (CD) and nuclear magnetic resonance (NMR) spectroscopy, were used to determine structural changes at FlhB post-autocleavage upon subunit binding and during the export specificity switch.

This combination of molecular microbiological, biochemical, and biophysical techniques was applied to answer the following outstanding questions in type III secretion:

1. What are the binding affinities of early subunit GRM interactions with FlhB? Do the affinities of these interactions impose order on early subunit export? If these affinities do impose order, this would be similar to late subunit export order.
2. Where do early subunits bind on $FlhB_C$ and what conformational changes may this cause at $FlhB_C$?
3. Is the early subunit binding site on $FlhB_C$ similar to sites found on early subunit export gates in other T3SS?
4. Why is FlhB autocleavage important to the export specificity switch?
5. What changes, if any, does FlhB experience to switch the export specificity of the flagellar type III export machinery?
6. How might FliK interact with FlhB to trigger the switch?

It is hoped that the results from this study provide a clearer picture of the early phase of flagellar subunit export and the transition between the early and late export phases. Early subunit binding to the export gate FlhB and the export specificity switch have great potential to help us understand how bacteria coordinate protein export to assemble intricate nanomachines. Efforts are already underway to engineer T3SSs to pump out biotechnologically relevant substrates. The GRM is short secretion signal that could easily be added to substrates to facilitate their export via the rapid fT3SS. Knowing the various binding affinities of early subunit GRM to FlhB could lead to more efficient substrate export. Finally, understanding how T3SSs operate in phases will open a pipeline for scientists to design rational, novel therapeutics against a wide variety of bacteria as we are faced with the rising challenge of antimicrobial resistance and persistence. These cases make it even more crucial that we understand how type III secretion systems impose order on subunit export.

Chapter 2

Materials and Methods

2.1 Reagents, buffers, and media

A full list of all reagents, buffers, and media used can be found in Table A1.1.

2.2 Enzymes

All enzymes were purchased from ThermoFisher Scientific and New England Biolabs and used according to manufacturer's instructions. Enzymes were stored at -20°C.

2.3 Oligonucleotides

All oligonucleotides used in this study were designed using Serial Cloner and synthesised by Sigma Aldrich at desalt purification and without modification. Primers were stored in deionised water (MilliQ) at -20°C. Table A1.2 provides a full list of primers used in this study.

2.4 Bacterial Strains

Table A1.3 lists the name, description, and source of all strains used in this study. Strains were grown in Luria-Bertani (LB) broth (Formedium) or agar (1.5% agar, Formedium) with appropriate antibiotic at 30°C or 37°C. Cells were harvested at various intervals during the growth cycle as measured by absorbance at 600nm (A_{600}) using a Cecil CE 2041 spectrophotometer. All strains were kept in 25% v/v glycerol at -80°C.

2.5 Plasmids

Plasmids used in this study can be found in Table A1.4. All plasmids were stored in deionised water (MilliQ) at -20°C.

2.6 DNA isolation, manipulation, and analysis

2.6.1 Isolation of DNA

Plasmids were isolated from 6 ml of overnight broth culture using the QIAprep spin miniprep kit (Qiagen) or PureLink quick plasmid mini kit (Life Technologies, ThermoFisher Scientific) following manufacturer's instructions. Plasmids were eluted

from supplied columns with 50 μl pre-warmed (50°C) deionised water for use in the polymerase chain reaction (PCR), restriction digests, or sequencing. DNA concentration and purity was measured using a nandrop ND1000 (Thermo Scientific).

Genomic DNA from an overnight culture of *Salmonella enterica* serovar Typhimurium SJW1103 was purified using the GeneJET genomic DNA purification kit (ThermoFisher Scientific) following manufacturer's instructions.

2.6.2 Analysis and extraction of DNA from agarose gels

Analysis of DNA was carried out using agarose gel electrophoresis with 1% w/v agarose (Sigma Aldrich) gels in 1x Tris-Acetic Acid-EDTA (TAE) buffer containing 5 μg ml^{-1} ethidium bromide. Samples were resuspended in gel loading buffer, purple (NEB) before being loaded into the gel. DNA was visualised using UV illumination and fragment size was estimated compared to a 2-log DNA ladder (NEB). QIAquick gel extraction kit (Qiagen) or GeneJET gel extraction kit (Life Technologies, ThermoFisher Scientific) were used to excise DNA from 1% w/v agarose gel in 1x TAE buffer following manufacturer's instructions. DNA was eluted from supplied columns with 50μl pre-warmed (50°C) deionised water for use in downstream applications.

2.7 Construction of bacterial strains

2.7.1 Scarless mutagenesis or deletion within the Salmonella chromosome

Salmonella enterica serovar Typhimurium strains were constructed using a modified method from the one described by Blank, Hensel, and Gerlach (Figure 2.1) (221). Briefly, oligonucleotides were designed to target the appropriate gene with replacement by a kanamycin resistance cassette containing a I-*Sce*I endonuclease recognition site (encoded on plasmid pWRG717). PCR using pWRG717 as template was performed to amplify the resistance cassette flanked on its 5' and 3' ends by 45 nucleotides homologous to the gene region of choice. The lambda red recombinase, controlled by heat shock-inducible promoter, was used to knock in the resistance cassette (pWRG730, $chloramphenicol^{R}$).

Salmonella enterica serovar Typhimurium SJW1103 was transformed with pWRG730 per the method outlined in 2.10.2. Cells were plated on LB agar with chloramphenicol

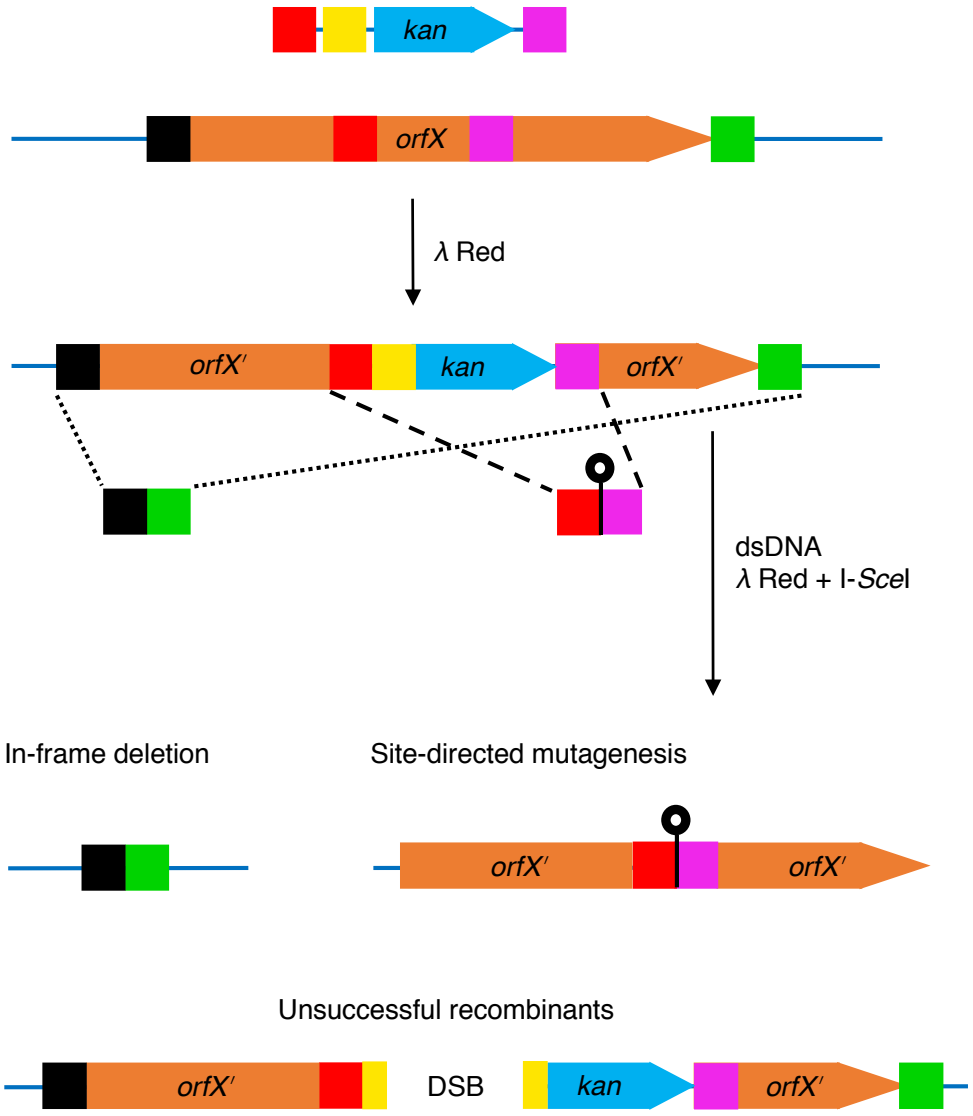


Figure 2.1 Scarless mutagenesis or in-frame deletion within the *Salmonella* chromosome using the lambda red (*λ* Red) recombinase

Briefly, a kanamycin resistance cassette (*kan*) and the I-*Sce*I recognition site (yellow) are amplified by PCR from pWRG717 plasmid using primers with 45 base pair homology extensions (red and pink). This construct is integrated into the *Salmonella* Typhimurium genome at *orfX* (any gene of interest) using *λ* Red-mediated recombination induced by incubating the cells at 42°C. A second round of *λ* Red-mediated recombination with synthetic double-stranded DNA (dsDNA, ~1000 base pairs in length) can be used to remove *orfX*, resulting in an in-frame deletion (black and green), or to mutate a specific region of *orfX* (red and pink). This results when the synthetic dsDNA with homologous regions recombines with the chromosome. Unsuccessful recombinants are selected against by the I-*Sce*I endonuclease induced with the addition of anhydrotetracycline to the growth medium, resulting in double stranded breaks (DSB). The plasmid encoding both the *λ* Red recombinase and I-*Sce*I endonuclease could be removed from the cell by incubation at 42°C without antibiotic. Figure adapted from Blank, Hensel, and Gerlach, 2011.

10 μg ml^{-1} and allowed to grow at 30°C overnight. A single colony was inoculated into 5 ml LB broth with chloramphenicol 10 μg ml^{-1} and incubated with 200 rpm shaking at 30°C overnight.

Overnight culture was diluted 1:100 into 100 ml LB broth with chloramphenicol 10 μg ml^{-1}. Cultures carrying the pWRG730 plasmid were incubated at 30°C with 200 rpm shaking until an A_{600} of 0.3. The lambda red recombinase was then induced by incubation at 42°C, 100 rpm, for 12.5 minutes. Cultures were chilled on ice for 15 minutes before being washed three times with deionised water and once with 10% v/v glycerol as described in 2.10.2. Strains were resuspended in 10% v/v glycerol used immediately or stored at -80°C for later use. Successful recombinant strains were selected by growth on LB agar containing chloramphenicol 10 μg ml^{-1} and kanamycin 25 μg ml^{-1}. Strains were confirmed by phenotype assays and sequencing. pWRG730 carrying the lambda red recombinase and I-*Sce*I endonuclease was cured from the cell by overnight incubation at 42°C. The loss of pWRG730 was confirmed by plating colonies on LB agar and LB agar with chloramphenicol 10 μg ml^{-1} and selecting only strains which grew exclusively on LB agar overnight at 37°C.

To remove the kanamycin resistance cassette, electrocompetent cells of kanamycin-resistance strains carrying active lambda red recombinase and tetracycline-inducible I-*Sce*I endonuclease on pWRG730 were transformed with a DNA insert produced synthetically (Sigma Aldrich) or by PCR. Cultures were incubated and the recombinase was induced as described above and electrocompetent cells prepped as in 2.10.2. PCR was used to generate a 1000 base pair DNA insert. This insert was homologous to the surrounding region of the appropriate gene. The insert could be used to knock out the gene of interest, leaving a scarless deletion, or to mutate the chromosomal copy of the gene depending on desired strain characteristics. Cells were plated on LB agar containing 100 ng ml^{-1} of anhydrotetracycline (AHT) to induce the expression of the I-*Sce*I endonuclease. The I-*Sce*I endonuclease cut the chromosome of any cells still carrying the kanamycin-resistance cassette, killing those cells and leaving only colonies of recombinant strains. Successful recombination was confirmed by phenotype assays and sequencing. pWRG730 carrying the lambda red recombinase and I-*Sce*I endonuclease was liberated by plating confirmed strains on

LB agar and incubating them overnight at 42°C. The loss of pWRG730 was confirmed by plating colonies on LB agar and LB agar with chloramphenicol 10 μg ml^{-1} and selecting only strains which grew exclusively on LB agar overnight at 37°C.

2.8 Polymerase Chain Reaction (PCR)

2.8.1 Amplification of DNA by PCR.

DNA of interest was amplified using Q5 hot start high-fidelity DNA polymerase (New England Biolabs) following manufacturer's recommendations in a MJ Mini Personal thermocycler (Bio-Rad) using thin-walled, flat-capped tubes. Q5 hot start high-fidelity DNA polymerase was chosen as it allows for room temperature setup, has 3'→5' exonuclease activity, and an error rate lower than *Pfu* or *Taq* polymerase. Extension time was adjusted based on the length of the target gene and T_m was adjusted based on the chemistry of the primer pair. *Salmonella enterica* serovar Typhimurium strain SJW1103 genomic DNA or previously constructed plasmid served as templates for the reactions. Reaction conditions and an example reaction are below.

Reaction mixture

Polymerase buffer (5X)	10 μl
10mM dNTPs	1 μl
10μM primer pair (forward and reverse)	5 μl
Template DNA (50ng)	1 μl
Q5 hot start high-fidelity polymerase (2U)	0.5 μl
Q5 GC enhancer (5X)	10 μl
Nuclease-free water	22.5 μl
Total reaction volume	50 μl

PCR reaction setup

1. Initial denaturation	98°C, 30 sec
2. Denaturation	98°C, 10 sec
3. Annealing	50-72°C, 30 sec
4. Extension	72°C, 20 sec kb^{-1}
5. Final Extension	72°C, 2 min
6. Hold	4°C, ∞

Steps 2-4 were cycled 25 times. QIAquick gel extraction kit (Qiagen) or GeneJET gel extraction kit (Life Technologies, ThermoFisher Scientific) were used to clean up PCR products following manufacturer's instructions.

2.8.2 Overlap extension PCR

Overlap extension PCR is often used to splice together DNA molecules or to introduce mutations into a target DNA molecule. Here, it was mainly used for the latter purpose. Overlap extension PCR was performed with the reagents and protocol described in 2.9.1 with slight modification. Two rounds of PCR were used, the products of the first round serving as the templates for the second round (Figure 2.2).

Primer sets were designed so that the mutation being introduced fell in the middle of each mutagenesis primer. In total, four primers were designed for each reaction. Primers A and D flanked the gene of interest. Primers B and C, reverse complementary to one another, included the desired mutation and flanking region. In the first round of PCR, two reactions were carried out. Primers A and B were mixed as a pair, as were primers C and D. Agarose gel electrophoresis was used to confirm that only a single PCR product of proper length was amplified. This product was then excised from the agarose gel and purified, removing any background template from the first reaction which could contaminate the second round of PCR.

In the second round of overlap extension PCR, primer pair A and D were used to amplify the overall gene region by mixing the two templates generated in the first round. Again, the PCR product was confirmed to be a single product of correct length using agarose gel electrophoresis and excised from the agarose gel. This final product, containing the desired mutation, was then used in the construction of recombinant plasmid or bacterial strains.

2.9 Construction of recombinant plasmids

2.9.1 Preparation of vector and insert DNA

Vector and insert DNA were digested with restriction enzymes (NEB) at 37°C for 2 hours in a MJ Mini Personal thermocycler (Bio-Rad) to generate compatible termini for ligation. Vector DNA was also incubated at 37°C for 2 hours with 1 U calf intestinal

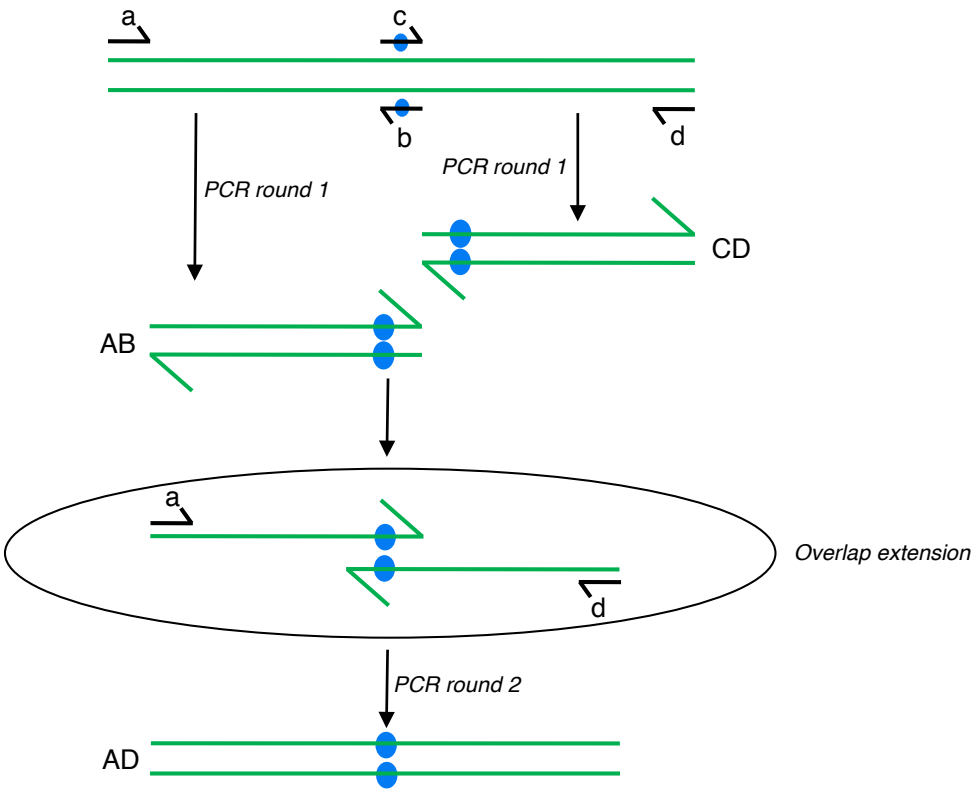


Figure 2.2 Overlap extension PCR

Overview of overlap extension PCR. The blue circles represent a point mutation, but overlap extension PCR can also be used to introduce short deletions or insertions as well.

alkaline phosphatase (NEB) to dephosphorylate the 5' end and prevent vector self-ligation. If insert DNA was not purified by extraction from a TAE agarose gel, 10 U of restriction enzyme DpnI was included in the restriction digest mixture to degrade any remaining methylated template DNA. Sample from both digestions was mixed with gel loading buffer, purple, and analysed for proper and complete digestion on a 1% w/v agarose gel in TAE buffer. Digest reactions were cleaned using the QIAquick gel extraction kit (Qiagen) or GeneJET gel extraction kit (Life Technologies, ThermoFisher Scientific) following manufacturer's instructions.

2.9.2 Ligation of vector and insert DNA

DNA concentration was calculated using a nanodrop ND1000 (Thermo Scientific). Vector and insert DNA were mixed in a 1:3 molar ratio with 1 U T4 DNA ligase (ThermoFisher Scientific) and reaction buffer containing adenosine triphosphate (ATP) and $MgCl_2$. The ligation reaction was incubated at 20°C for 1 hour in a MJ Mini Personal thermocycler (Bio-Rad). The T4 DNA ligase was inactivated by heating the reaction mixture to 70°C for 5 minutes.

2.9.3 Confirmation of recombinant plasmid

The ligation mixture (3.5 μl) was transformed into 45 μl XL-1 blue or DH10β cells in 10% v/v glycerol. Colonies appearing after overnight incubation at 37°C on LB agar with appropriate antibiotic were selected for colony PCR using Reddymix PCR Master Mix 2x (Thermo Scientific) with a primer pair specific to the vector and desired insert. Standard reaction conditions as recommended by the manufacturer were used, the T_m and extension time adjusted per the primer pair and length of insert. Plasmid was prepped as above from strains carrying the correct sized insert on the recombinant plasmid and sent to Source Bioscience or the University of Cambridge Department of Biochemistry sequencing facility for confirmation by Sanger sequencing. DNA sequence alignment was performed using the Multalin online tool (Corpet, F. 1988). Glycerol stocks with a final concentration of 25% v/v glycerol were made for each strain carrying a confirmed recombinant plasmid and stored at -80°C.

2.10 Preparation and transformation of electrocompetent cells

2.10.1 Escherichia coli strains C41 (DE3), XL-1 blue, and DH10β

Overnight culture of *E. coli* strains was diluted 1:100 in 500 ml of sterile LB broth. Culture was incubated with 200 rpm shaking at 37°C to an A_{600} of 0.5. Cells were chilled on ice for 15 minutes before being collected by centrifugation at low speed (3000 xg, 10 minutes) to minimise cell lysis. Cells were resuspended twice in 250ml chilled deionised water and then collected by centrifugation. A final resuspension and wash in 25 ml of chilled deionised water was performed. Cells were collected by centrifugation and resuspended in 25 ml of chilled sterile 10% v/v glycerol before a final centrifugation step. Prepared cells were resuspended to a final volume of 2 ml of chilled sterile 10% v/v glycerol and aliquoted in 100 μl into 0.5 ml tubes (Starlab). Cells were immediately stored at -80°C.

Electrocompetent cells were thawed on ice before use. Cells (45 μl) were incubated on ice with 1 μl recombinant plasmid for 1 minute in a 1 mm cuvette (VWR). Cells were transformed by electroporation with the following settings: 1.25 kV, 200 Ω, 25 μF. Cells were quickly resuspended in 500 μl super optimal broth with catabolite repression (SOC) and recovered at 37°C for 1 hour with 200 rpm shaking. Transformed cells were then selected for by plating on LB agar with appropriate antibiotics. Cells containing recombinant plasmids encoding ampicillin resistance were selected for on LB agar with carbenicillin to reduce the prevalence of satellite colonies.

2.10.2 Salmonella enterica serovar Typhimurium SJW1103 and JR501

Overnight culture of *Salmonella* strains was diluted 1:100 in 500 ml of sterile LB broth. Culture was incubated with 200 rpm shaking at 37°C to an A_{600} of 0.5. Cells were chilled on ice for 15 minutes before being collected by centrifugation at low speed (3000 xg, 10 minutes) to minimise cell lysis. Cells were resuspended twice in 250 ml chilled deionised water and incubated on ice for 30 minutes before collection by centrifugation between washes. A final resuspension and wash in 25 ml of chilled deionised water was performed. Cells were collected by centrifugation after incubation on ice for 30 minutes and resuspended in 25 ml of chilled sterile 10% v/v glycerol. Cells were immediately collected by centrifugation. Prepared cells were resuspended

to a final volume of 2 ml of chilled sterile 10% v/v glycerol and aliquoted in 100 μl into 0.5 ml tubes (Starlab). Cells were immediately stored at -80°C.

Electrocompetent cells were thawed on ice before use. Cells (45 μl) were incubated on ice with 1 μl recombinant plasmid for 1 minute in a 1 mm cuvette (VWR). Cells were transformed by electroporation with the following settings: 1.25 kV, 200 Ω, 25 μF. Cells were quickly resuspended in 500 μl (SOC) and recovered at 37°C for 1 hour with 200 rpm shaking. Transformed cells were then selected for by plating on LB agar with appropriate antibiotics. Cells containing recombinant plasmids encoding ampicillin resistance were selected for on LB agar with carbenicillin to reduce the prevalence of satellite colonies.

Before transformation in SJW1103, recombinant plasmids were transformed into and purified from JR501 to allow for higher transformation efficiency.

2.11 Expression of recombinant protein

2.11.1 Expression of intein-chitin binding domain fusion protein

E. coli C41 (DE3) cells were transformed with recombinant expression plasmid DNA and plated on LB agar containing appropriate antibiotic. LB broth (1 L) with appropriate antibiotic was inoculated with colonies and incubated at 37°C with 200 rpm shaking until an A_{600} of 0.7. Protein expression under the control of the T7 promoter was induced with the addition of isopropyl β-D-1-thiogalactopyranoside (IPTG, Generon) to a final concentration of 1 mM. Cells were incubated at 30°C with 200 rpm shaking for a further 4-5 hours.

2.11.2 Expression of His-tagged proteins

E. coli C41 (DE3) cells were transformed with recombinant expression plasmid DNA and plated on LB agar containing appropriate antibiotic. 2x Yeast Tryptone broth (1 L) (2xYT, Formedium) with appropriate antibiotic was inoculated with colonies and incubated at 37°C with 200 rpm shaking until an A_{600} of 0.7. Protein expression under the control of the T7 promoter was induced with the addition of IPTG to a final concentration of 1 mM. Cells were incubated at 30°C with 200 rpm shaking for a further 4-5 hours.

2.11.3 Expression of proteins in M9 minimal medium

E. coli C41 (DE3) cells were transformed with recombinant expression plasmid DNA and plated on LB agar containing appropriate antibiotic. Plates were incubated overnight at 37°C. Ammonium chloride was used as the nitrogen source and D-glucose was used as the carbon source.

2.11.3.1 M9 minimal medium containing ^{15}N-ammonium chloride and/or ^{13}C-labeled glucose

M9 minimal broth (10 ml) containing appropriate antibiotic was inoculated with 3-4 colonies overnight at 37°C. Multiple colonies were used to increase cell growth in the minimal medium. Overnight cultures were diluted 1:100 in 500 ml M9 minimal broth containing appropriate antibiotic and $^{15}NH_4Cl$, 98 atom %, and/or D-glucose-$^{13}C_6$, 99 atom % (Sigma). Cultures were incubated at 37°C with 200 rpm shaking until an A_{600} of 0.7. Protein expression under the control of the T7 promoter was induced with the addition of IPTG to a final concentration of 1 mM. Cells were incubated at 20°C with 200 rpm shaking for a further 16 hours.

2.11.3.2 M9 minimal medium containing ^{15}N-ammonium chloride, deuterated ^{13}C-labeled glucose, and 99% D_2O

M9 minimal broth (10 ml) containing appropriate antibiotic was inoculated with 3-4 colonies overnight at 37°C. Cells were passaged via a 1:20 dilution in 10 ml of M9 minimal broth containing appropriate antibiotic and increasing amounts of D_2O (0%, 50%, 70%, 90%, 90%, 99%, 99%, 99%). Each passage was incubated overnight at 37°C with 200 rpm shaking. Final overnight cultures were diluted 1:100 in 500 ml M9 minimal broth containing appropriate antibiotic and $^{15}NH_4Cl$, 98 atom %, D-Glucose-$^{13}C_6$,1,2,3,4,5,6,6-d_7, 99 atom %, and 99% D_2O (Sigma). Cultures were incubated at 37°C with 200 rpm shaking until an A_{600} of 0.7. Protein expression under the control of the T7 promoter was induced with the addition of IPTG to a final concentration of 1 mM. Cells were incubated at 20°C with 200 rpm shaking for a further 16 hours.

2.12 Protein purification

2.12.1 Purification of $FlhB_C$-intein-chitin binding domain fusion protein

$FlhB_C$ residues 219-383 G383A fused to the intein-chitin binding domain ($FlhB_C$-inteinCBD) was used in isothermal titration calorimetry, crystallography, and nuclear magnetic resonance spectroscopy experiments. Purification of $FlhB_C$-inteinCBD was performed as previously described with slight modifications (222, 223). InteinCBD was chosen as an expression and purification partner because prior studies successfully purified FlhB using this system and for its cost-effectiveness and ease of use. Briefly, cells were resuspended in 25 ml buffer A with 2 U DNaseI (ThermoFisher Scientific) and 1 tablet of cOmplete protease inhibitor cocktail, EDTA-free (Roche). The resuspension was homogenised using a glass tissue homogeniser (VWR). Cells were lysed by cell disruption at 30 kpsi (Constant Systems Ltd). The soluble fraction was collected by centrifugation at 100,000 xg, 4°C for 1 hour and filtered through a 0.2 µm Minisart hydrophilic cellulose acetate filter (Sartorius). The soluble fraction was then bound to pre-equilibrated chitin resin (NEB) and washed with 10 column volumes (CV) buffer A. $FlhB_C$ was cleaved from the resin-bound inteinCBD tag by incubating the proteins with buffer A supplemented with 100 mM dithiothreitol for 16 hours at 4°C (DTT, Melford Laboratories).

Eluate containing $FlhB_C$ was dialysed against 5 L buffer B (50 mM Tris-HCl, pH 7.4, 150 mM NaCl) for two hours using a 5 kDa cutoff regenerated cellulose dialysis membrane tubing (Spectrum Labs). Purification continued using an Äkta Express system (GE Healthcare) and a pre-equilibrated Hitrap SP HP 5 ml cation exchange column (GE Healthcare). The column with bound protein was washed with 10 CV buffer B. Protein was eluted using a gradient of up to 1.2 M NaCl. Fractions with $FlhB_C$ were determined by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) stained with Coomassie R250 and were collected for further purification using a Superdex 75 size exclusion column (GE Healthcare) pre-equilibrated with buffer B. The final $FlhB_C$ concentration was measured using a DC Protein Assay (Bio-Rad) with known Bovine Serum Albumin (BSA, GE Healthcare) standards. Protein was stored at 4°C or -80°C depending on application. If stored at -80°C, protein was frozen immediately in liquid nitrogen.

2.12.2 Purification of intein-chitin binding domain tagged proteins

$FlhB_C$ residues 211-383 G383A and variants fused to the intein-chitin binding domain ($FlhB_{C211}$-inteinCBD) were used in circular dichroism scanning and thermal melt experiments. Cells were resuspended in 25 ml buffer A with 2 U DNaseI and 1 tablet of cOmplete protease inhibitor cocktail, EDTA-free. The resuspension was homogenised using a glass tissue homogeniser. Cells were lysed by cell disruption at 30 kpsi. The soluble fraction was collected by centrifugation at 60,000 xg, 4°C for 1 hour. The soluble fraction was then bound to pre-equilibrated chitin resin and washed with 10 CV buffer A. $FlhB_C$ variants were cleaved from the resin-bound inteinCBD tag by incubating the proteins with buffer A supplemented with 100 mM DTT for 16 hours at 4°C.

Eluate containing $FlhB_C$ was dialysed twice against 5 L buffer C for 8 hours at 4°C using fresh 3.5 kDa cutoff SnakeSkin dialysis membrane tubing (ThermoFisher Scientific) each time. $FlhB_{C211}$ purity was determined by SDS-PAGE stained with Instant Blue. If protein was determined to be less than 95% of the fraction, $FlhB_{C211}$ was further polished using a pre-equilibrated Superdex 75 size exclusion column (GE Healthcare) with buffer C. The final $FlhB_{C211}$ concentration was measured using a nanodrop ND1000 set to 280 nm, an extinction coefficient of 22460 M^{-1} cm^{-1}, and a molecular mass of 19651.66 g/mol. Protein was stored at 4°C until use.

2.12.3 Purification of HexaHis-tagged proteins

Cells expressing HexaHis-tagged proteins were resuspended in 25 ml buffer D with 2 U DNaseI and 1 tablet of cOmplete protease inhibitor cocktail, EDTA-free. The resuspension was homogenised using a glass tissue homogeniser. Cells were lysed by cell disruption at 30 kpsi. The soluble fraction was collected by centrifugation at 60,000 xg, 4°C for 1 hour. The soluble fraction was then bound to a pre-equilibrated HisTrap Excel 1 ml column (GE healthcare) attached to a BioLogic DuoFlow fast protein liquid chromatography (FPLC) machine (Bio-Rad) and washed with 10 CV buffer D. HexaHis-tagged proteins were eluted using buffer E. Fractions were analysed by SDS-PAGE stained with Instant Blue and those fractions containing the hexaHis-tagged protein of interest were pooled. HexaHis-tagged proteins were dialysed against buffer conditions required for the relevant experiment using fresh 3.5

kDa cutoff SnakeSkin dialysis membrane tubing. Dialysis was carried out overnight at 4°C to ensure removal of imidazole. Final protein concentration was measured using a DC Protein Assay against known BSA standards. Protein was used immediately or stored at 4°C until use.

2.12.4 Purification of co-expressed FlgB and FlgC

Recombinant pTXBI plasmid containing HexaHis-3xFLAG-FlgB and FlgC was expressed in *E. coli* C41 cells at 20°C overnight with 180 rpm shaking. This condition allowed for optimal expression of both proteins. Cells were collected by centrifugation and resuspended in 25 ml of buffer F with 2 U DNaseI and 1 tablet of cOmplete protease inhibitor cocktail, EDTA-free. The resuspension was homogenised using a glass tissue homogeniser. Cells were lysed by cell disruption at 30 kpsi. The soluble fraction was collected by centrifugation at 60,000 xg, 4°C for 1 hour. The soluble fraction was then bound to a pre-equilibrated HisTrap Excel 1 ml column (GE healthcare) attached to a BioLogic DuoFlow fast protein liquid chromatography (FPLC) machine (Bio-Rad) at a low flow rate of 0.5 ml min^{-1}. The column was then washed with 10 CV buffer F at the same low flow rate. Proteins were eluted using buffer F supplemented with 300 mM imidazole in 1 ml fractions. Fractions were analysed by immunoblotting (Section 2.14) using mouse anti-FLAG M2 (1:5000) and mouse anti-CBD (1:2000) for co-elution of FlgB and FlgC.

2.13 Electrophoresis and staining of proteins

Proteins were analysed by size separation on SDS polyacrylamide gels using the Laemmli method and were visualised using Coomassie Brilliant Blue R250 (Sigma Aldrich) or Instant Blue (Expedeon). Samples were resuspended in SDS loading buffer containing 140 mM β-mercaptoethanol (βME, Sigma Aldrich) and vortexed vigorously to remove any aggregates. Samples were then boiled at 95°C for 7 minutes to denature the proteins. Once cooled, samples were loaded onto 12.5-15% v/v polyacrylamide (Bio-Rad or Severn Biotech) resolving gels based on the predicted molecular weight of the protein(s) of interest. Stacking gels were 10% v/v polyacrylamide. After electrophoresis, usually 140 V for 65 minutes, gels were stained in Instant Blue and destained in deionised water or stained and fixed with 0.25% Coomassie Brilliant Blue R250, 10% acetic acid (Fisher Scientific), and 50% methanol

(Sigma Aldrich) for 10 minutes before destaining in 20% methanol and 7% acetic acid for an hour.

2.14 Immunoblotting of proteins

Samples separated by SDS-PAGE were transferred onto nitrocellulose membrane, 0.45 μm pore size (GE Healthcare), at 400 mA and 100 V for one hour in transfer buffer in a TE22 wet transfer tank (Hoefer). After transfer, membrane was blocked overnight at 4°C with 20 ml phosphate buffered saline (Hutchison Media Unit, University of Cambridge) with 0.05% v/v Triton X-100 (VWR) (PBST) and 5% w/v skim milk power (Marvel). Blots were then washed one time in 10 ml PBST to remove the skim milk.

Fresh blocking solution (10 ml) with diluted primary antibody was added to the blot and allowed to incubate at room temperature for one hour. Primary antibody depended upon the protein sample being probed. A full list of antibodies and dilutions used in this study can be found in Table AI.5. Where FlhB or $FlhB_{CT3xFLAG}$ was being detected, 0.5 ml of *SJW1103Δ flhB* soluble lysate (in PBS) was added to the blocking solution with primary antibody. The addition of the *flhB* null lysate greatly reduced the appearance of contaminating bands when analysing samples.

After incubation with primary antibody, blots were washed three times, 5 minutes each in fresh blocking solution. Fresh blocking solution with 1:10000 IRDye 680RD donkey anti-mouse and/or IRDye 800CW donkey anti-rabbit infrared fluorescent secondary antibodies (Li-Cor) was then added and incubated at room temperature in the dark for one hour. Blots were washed twice, 5 minutes each in PBST before being washed twice, 5 minutes each in phosphate buffered saline (PBS). Blots were analysed with the Li-Cor Odyssey CLx system set to scan on both available infrared channels (680 nm and 800 nm).

2.15 Analysis of protein concentration

2.15.1 Bio-Rad DC Protein assay

Protein was diluted 1:10 in buffer. The same buffer served as a buffer absorbance control. BSA standards (0.25, 0.5, 1.0, and 2.0 mg ml^{-1}) were stored in 0.9% w/v saline

solution at -20°C. They were thawed at room temperature as needed. Deionised water was used as the absorbance blank. All samples were performed in duplicate.

Sample (20 μl of protein, buffer, or BSA standard) was added to a 10x4x45 mm polystyrene cuvette (Sarstedt). Reagent A (100 μl) was added to each sample, followed by 800 μl of Reagent B. Samples were vortexed and incubated for 10 minutes in the dark. Sample absorbance was measured at 750 nm, or A_{750}, with a spectrophotometer (Cecil). Any A_{750} from buffer was subtracted from the protein sample. The unknown protein sample was compared to the BSA standards using a best fit line to determine protein concentration. Protein purity was confirmed using SDS-PAGE and concentration was reassessed if needed.

2.15.2 Nanodrop spectrophotometer

Protein was diluted 1:10 in buffer. The same buffer served as an absorbance blank control. Sample (1 μl) was analysed by a scan from 200-400 nm. Concentration was determined by the absorbance at 280nm, or A_{280}, using the protein's predicted molecular weight and extinction coefficient. Protein purity was confirmed using SDS-PAGE and concentration was reassessed if needed.

2.16 Subunit export assay

Export assays were performed as previously described with few modifications (86). *S. enterica* serovar Typhimurium strains were first transformed with relevant pBAD18amp plasmids as described above. The A_{600} of overnight cultures was measured and culture was diluted in 10 ml LB with appropriate antibiotics and 0.02% w/v L-arabinose (Acros Organics) to an A_{600} of 0.05. Cultures were incubated at 37°C with 200 rpm shaking until the culture reached an A_{600} of 0.2. Cultures were then washed and transferred to 10 ml fresh LB with appropriate antibiotics and 0.02% w/v L-arabinose. A 1 ml sample was taken and centrifuged at 17,000 xg, 4°C. The cell pellet was saved as the "whole cell" fraction. Supernatant was recovered and passed through a 0.2 μm Minisart hydrophilic cellulose acetate filter. Exported subunits were precipitated from filtered supernatant with 10% v/v trichloroacetic acid (TCA) and 1% v/v Triton X-100 overnight at 4°C. Precipitated protein was centrifuged at 17,000 xg, 4°C for 20 minutes. Protein pellets were resuspended in 1 ml ice cold 100% acetone

(Sigma Aldrich) and vortexed vigorously to remove any remaining TCA. Precipiated protein was pelleted by centrifugation at 17,000 xg, 4°C for 20 minutes. Remaining acetone was allowed to evaporate. Whole cell and supernatant samples were resuspended in SDS-PAGE loading buffer, standardised against culture A_{600} with a 1:100 dilution for whole cell fractions and 1:50 for supernatant fractions. These were called time point "0" samples.

The remaining 9 ml of cultures were incubated for one hour at 37°C with 200 rpm shaking. A 1 ml sample of culture was taken and culture density was recorded at A_{600}. The sample was centrifuged at 17000 xg, 4°C for 5 minutes to pellet the cells. Whole cell and supernatant samples at 60 minutes, time point "60", were treated the same way as time point "0" samples. Whole cell and supernatant samples were resuspended in SDS-PAGE loading buffer, volume standardized against culture A_{600}. All samples were visualised with immunoblotting.

2.17 Swimming motility assay (in null and wildtype strains)

S. enterica serovar Typhimurium SJW1103 strains were transformed with pBAD18amp plasmids, if needed. The A_{600} of overnight cultures was measured and culture was diluted in 5 ml LB with appropriate antibiotics and 0.02% w/v L-arabinose to an A_{600} of 0.05. Cells were grown in LB broth with ampicillin 100 µg ml^{-1} at 37°C until the cultures reached an A_{600} of 1.0. Culture (2 µl) was then spotted into a pre-warmed motility agar plate with L-arabinose (0.002%- 0.2% w/v) to induce protein expression under the control of the P_{BAD} promoter. The control motility plate did not contain L-arabinose. Motility agar plates with SJW1103 and SJW1103*ΔflgE::Kan*^{R}, and SJW1103$flhB_{Ct3xFLAG}$ variants were incubated for 4 hours at 37°C. Motility plates with SJW1103*ΔflgC* were incubated for 18 hours at 30°C. Motility plates with SJW1103*ΔfliK* $flhB_{Ct3xFLAG}$ variants were incubated for 20 hours at 37°C.

2.18 Design of synthetic gate recognition motif (GRM) peptides

Synthetic peptides were ordered from Covalab (France). They were designed to be as hydrophilic as possible given the nature of the hydrophobic GRM to ensure they would resuspend well in a water-based, Tris buffer. All peptides in this study were N-terminally acetylated and C-terminally amidated to prevent peptides from ligating

together in solution. Peptides were synthesised to greater than 95% purity and confirmed by the manufacturer using HPLC and mass spectrometry. Upon receipt, peptides were resuspended in deionised water (MilliQ) to a final concentration of 5 mM for later dilution into buffer and stored at -20°C. A full list of peptides can be found in Chapter 3.

2.19 Isothermal titration calorimetry (ITC)

ITC was performed at 25°C using a VP-ITC system and all ITC data were analysed using the provided Origin software (MicroCal, GE Healthcare). The reference cell was filled with water. $FlhB_C$ was purified as previously described and concentrated using a VivaSpin 20 Concentrator with a 5k molecular weight cutoff (MWCO) (Vivascience, Sartorius) to 50 µM in 1.4 ml of buffer B. Protein was added to the sample cell and mixed, throughout the experiment, at 300 rpm. GRM peptide samples were diluted to 1 mM in buffer B and added to the syringe. All samples were degassed and heated to 25°C before use. Peptide (2 µl) was initially injected into the sample cell. The cell was allowed to equilibrate for 150 s. Peptide (6 µl) was then injected into the sample cell in 250 second intervals for a final ligand (peptide) injection volume of 176 µl (1 injection at 2 µl, 29 injections at 6 µl). The heat generated by the addition of peptide to $FlhB_C$ was recorded as was the energy needed to maintain a constant temperature in the sample cell as compared to the reference cell. The heat of dilution generated by the addition of peptide into buffer B was subtracted from the heats generated by the addition of peptide to $FlhB_C$. The integral of the calorimetric signal was calculated and a best fit line was used to determine the K_d of peptide to $FlhB_C$.

2.20 Crystallography of GRM peptides and FlhB

Purified $FlhB_C$ was crystallised in the presence of a fivefold molar excess of GRM peptides using the sitting drop vapor diffusion method. The full list of the screen conditions can be found in Figure 2.3 and Figure 2.4. Reservoir liquid (1 µl) was added to 1 µl protein:peptide mixture in a 24-well plate. The plate was sealed and incubated at 20°C. Crystals were checked 1 day, 3 days, 1 week, and 2 weeks after initial incubation. In preparation for x-ray diffraction, crystals were removed from the drops using a microloop and stereomicroscope. As all screen conditions contained a

Well ID	Sodium cacodylate	Zinc acetate	Propylene glycol	Well ID	Sodium cacodylate	Zinc acetate	Propylene glycol
1	pH 4.7 0.1 M	0.15 M	25%	25	pH 5.3 0.1 M	0.18 M	20%
2	pH 4.7 0.1 M	0.16 M	25%	26	pH 5.3 0.1 M	0.18 M	25%
3	pH 4.7 0.1 M	0.17 M	25%	27	pH 5.3 0.1 M	0.19 M	20%
4	pH 4.7 0.1 M	0.18 M	25%	28	pH 5.3 0.1 M	0.19 M	25%
5	pH 4.7 0.1 M	0.19 M	25%	29	pH 5.3 0.1 M	0.20 M	20%
6	pH 4.7 0.1 M	0.20 M	25%	30	pH 5.3 0.1 M	0.20 M	25%
7	pH 5.0 0.1 M	0.15 M	25%	31	pH 5.3 0.1 M	0.21 M	20%
8	pH 5.0 0.1 M	0.16 M	25%	32	pH 5.3 0.1 M	0.21 M	25%
9	pH 5.0 0.1 M	0.17 M	25%	33	pH 5.3 0.1 M	0.22 M	20%
10	pH 5.0 0.1 M	0.18 M	25%	34	pH 5.3 0.1 M	0.22 M	25%
11	pH 5.0 0.1 M	0.19 M	25%	35	pH 5.3 0.1 M	0.23 M	20%
12	pH 5.0 0.1 M	0.20 M	25%	36	pH 5.3 0.1 M	0.23 M	25%
13	pH 5.3 0.1 M	0.15 M	25%	37	pH 5.6 0.1 M	0.18 M	20%
14	pH 5.3 0.1 M	0.16 M	25%	38	pH 5.6 0.1 M	0.18 M	25%
15	pH 5.3 0.1 M	0.17 M	25%	39	pH 5.6 0.1 M	0.19 M	20%
16	pH 5.3 0.1 M	0.18 M	25%	40	pH 5.6 0.1 M	0.19 M	25%
17	pH 5.3 0.1 M	0.19 M	25%	41	pH 5.6 0.1 M	0.20 M	20%
18	pH 5.3 0.1 M	0.20 M	25%	42	pH 5.6 0.1 M	0.20 M	25%
19	pH 5.6 0.1 M	0.15 M	25%	43	pH 5.6 0.1 M	0.21 M	20%
20	pH 5.6 0.1 M	0.16 M	25%	44	pH 5.6 0.1 M	0.21 M	25%
21	pH 5.6 0.1 M	0.17 M	25%	45	pH 5.6 0.1 M	0.22 M	20%
22	pH 5.6 0.1 M	0.18 M	25%	46	pH 5.6 0.1 M	0.22 M	25%
23	pH 5.6 0.1 M	0.19 M	25%	47	pH 5.6 0.1 M	0.23 M	20%
24	pH 5.6 0.1 M	0.20 M	25%	48	pH 5.6 0.1 M	0.23 M	25%

Figure 2.3 Crystal screens used to co-crystallise $FlhB_C$ and GRM peptides

Chemical composition of two crystal screens for co-crystallisation of GRM peptides with $FlhB_C$. Wells 1-24 are screen GT1 and wells 25-48 are screen GT2.

Well ID	Sodium cacodylate	Zinc acetate	Propylene glycol
49	pH 5.5 0.1 M	0.19 M	25%
50	pH 5.5 0.1 M	0.19 M	28%
51	pH 5.5 0.1 M	0.20 M	25%
52	pH 5.5 0.1 M	0.20 M	28%
53	pH 5.5 0.1 M	0.21 M	25%
54	pH 5.5 0.1 M	0.21 M	28%
55	pH 5.6 0.1 M	0.19 M	25%
56	pH 5.6 0.1 M	0.19 M	28%
57	pH 5.6 0.1 M	0.20 M	25%
58	pH 5.6 0.1 M	0.20 M	28%
59	pH 5.6 0.1 M	0.21 M	25%
60	pH 5.6 0.1 M	0.21 M	28%
61	pH 5.7 0.1 M	0.19 M	25%
62	pH 5.7 0.1 M	0.19 M	28%
63	pH 5.7 0.1 M	0.20 M	25%
64	pH 5.7 0.1 M	0.20 M	28%
65	pH 5.7 0.1 M	0.21 M	25%
66	pH 5.7 0.1 M	0.21 M	28%
67	pH 5.8 0.1 M	0.19 M	25%
68	pH 5.8 0.1 M	0.19 M	28%
69	pH 5.8 0.1 M	0.20 M	25%
70	pH 5.8 0.1 M	0.20 M	28%
71	pH 5.8 0.1 M	0.21 M	25%
72	pH 5.8 0.1 M	0.21 M	28%

Figure 2.4 Crystal screens used to co-crystallise $FlhB_C$ and GRM peptides

Chemical composition of additional crystal screen for co-crystallisation of GRM peptides with $FlhB_C$. Wells 49-72 are screen GT3.

cryoprotectant (propylene glycol), crystals were immediately frozen and stored in liquid nitrogen.

$FlhB_C$ crystals were analysed using x-ray diffraction and resulting data was processed by Dr. Philip Hinchliffe. Diffraction data for $FlhB_C$-peptide crystals were collected on beamline I24 at UK Diamond Light Source using a Pilatus 6M detector. Data were collected with an oscillation angle of 0.2°, an exposure time of 0.3 seconds, and a crystal-detector distance of 579.73 mm with resolution set at 2.8 Å. Resulting data were indexed and integrated with MOSFLM (224), and scaled with Aimless in the CCP4 suite (225). Molecular replacement was performed using Phaser (226) and structure refined using Refmac (227).

2.21 Nuclear Magnetic Resonance (NMR) spectroscopy

NMR spectroscopy experiments were performed on ^{15}N labelled $FlhB_C$ and $^{15}N,^{13}C$ labelled $FlhB_C$ in buffer B purified as previously described. Sample was concentrated using a VivaSpin 20 Concentrator with a 5k MWCO to 200-300 μM in 450 μl and supplemented with a final concentration of 10% v/v deuterium oxide (D_2O). 2D [^1H, ^{15}N]-TROSY NMR experiments were recorded on a Bruker AV800 spectrometer (^1H frequency 800 MHz) equipped with a 5 mm HFCN/z cryoprobe while 3D ^1H, ^{15}N, ^{13}C triple resonance experiments were recorded on a Bruker AV600 spectrometer (^1H frequency 600 MHz) equipped with a 5 mm HFCN/z cryoprobe. Both spectrometers are part of the biomolecular NMR facility in the Department of Biochemistry. All NMR experiments were recorded at 25°C. All data were recorded and processed by Dr. Daniel Nietlispach, who also assisted with the analysis. Raw data were processed using the Azara software package (W. Boucher, unpublished). Sequential backbone assignments were obtained using the CcpNMR analysis software based on the following NMR experiments: 3D HNCA, 3D HN(CO)CA, 3D HN(CA)CB, 3D HN(COCA)CB, and 3D ^{15}N NOESY (228).

Titration experiments with labelled $FlhB_C$ at a concentration of 300 μM were conducted by addition of GRM peptide at an initial concentration of 4.8 mM in buffer B. Peptide was added in 20 μl increments until 2D [^1H, ^{15}N]-TROSY spectra showed no further changes. All samples were monitored by 2D Selective Optimized-Flip-Angle Short-

Transient Heteronuclear Multiple Quantum Correlation (SOFAST-HMQC) experiments with a receiver gain of 128 and a total of 8 scans. Spectra at GRM peptide saturation were recorded and compared to the unbound $FlhB_C$ reference spectra to determine the different chemical shifts upon a subunit-binding event. The changes in chemical shifts (Δδ) were calculated by subtracting the peptide-bound shifts by the unbound shifts to yield $\Delta\delta_H$ and $\Delta\delta_N$. Δδ was found using the following formula: $\Delta\delta = \sqrt{((\Delta\delta_H)^2+((\Delta\delta_N)^2/6))}$ (229). Δδ was plotted in comparison to the $FlhB_C$ residue number.

2.22 Bioinformatics

2.22.1 Analysis of protein structure using PyMol

Protein structures were downloaded from the Research Collaboratory for Structural Bioinformatics Protein Database (RCSB PDB) as .pdb files. $FlhB_C$ and $FliK_C$ from *S. enterica* serovar Typhimurium have the accession number 3B0Z and 2RRL, respectively (230). Structures were loaded into PyMol v1.5.0.5 Enhanced for Mac OS X. Where necessary, the PyMol measurement wizard was used to measure the distance between atoms and the PyMol mutagenesis wizard was used to mutate selected residues.

2.22.2 Alignment of protein sequences

Protein sequences were downloaded from UniProt (231). Sequences were aligned using the Clustal Omega multiple sequence alignment tool hosted by the European Bioinformatics Institute (232). Default parameters were used, except output of sequences was changed to “as aligned”.

2.23 Circular dichroism (CD) of proteins

All CD measurements were carried out with an AVIV 410 CD instrument (AVIV Biomedical) at the Biophysics Facility in the Department of Biochemistry of the University of Cambridge. Pure nitrogen gas was used to purge the xenon lamp of oxygen (below 5 parts per million). $FlhB_{C211}$ and its variants were purified and dialysed into buffer C as previously described. Protein was concentrated to 0.1-0.2 mg ml^{-1} using a VivaSpin 20 Concentrator with a 5k MWCO. 400 μl of protein was added to a 1 mm glass cuvette. All data were collected using AVIV 410 software. Data were

analysed using Kaleidagraph (Synergy Software). All CD experiments were planned, conducted, and analysed with Dr. Katie Kemplen.

2.23.1 General CD scans

Protein was scanned at 25°C. Each protein was scanned 3 times and the average taken. Scans were conducted in a range from 260 nm to 200 nm, in 1 nm intervals with 3 seconds averaging time at each interval. A reference scan using buffer was subtracted from the average protein scan. The resulting averaged protein curve was smoothed using standard settings in the AVIV 410 software. The mean residue ellipticity (MRE) was calculated using the following equation:

$$MRE = \frac{mdegree}{Molar\ concentration * path\ length * \#\ of\ amino\ acids}$$

The averaged mdegree (mdeg) was used. Molar concentration was determined using methods described previously. Path length was set to 0.1 cm and each protein contained 173 residues, except for $FlhB_{C211}$ $W_{353}T$, which only had 143 residues.

2.23.2 Analysis of protein melting temperature by CD

Protein was scanned in a range from 25°C to 95°C to 25°C, in 1°C intervals with 3 seconds averaging time at each interval. Protein was held at 95°C for 5 minutes before the scan from 95°C to 25°C was completed. All measurements were recorded at 222 nm.

To find the best fit line to the data, the following equation was used (233, 234):

$$y = m3 + (m5 \times m0) + (m4 - m3 + \big((m6 - m5) \times m0\big))/\ (1 + \exp\ (0.12 \times (m7 \times \ln\ (\left(\frac{1}{m2}\right) \times m0)) - (\left(\left(\frac{1}{m0}\right) - \left(\frac{1}{m2}\right)\right) \times (m1 - \left(\left(\frac{1}{m2}\right)\right) \times m7))))$$

The terms for the equation are defined below.

m0	x variable
m1	∆H (change in enthalpy) (Greenfield, 2007; Kemmer and Keller, 2010)
m2	melting temperature (in Kelvin)
m3	y-intercept for folded/native baseline

m4	y-intercept for the unfolded baseline
m5	slope of the native/folded baseline
m6	slope of the unfolded baseline
m7	ΔCP (change in heat capacity)

2.24 Whole cell samples of FlhB variants with and without FliK

Salmonella enterica serovar Typhimurium SJW1103 strains carrying a chromosomal copy of $flhB_{Ct3xFLAG}$ and its variants were plated on LB agar to isolate individual colonies. Strains with a kanamycin resistance cassette in place of *fliK* were plated on LB agar with kanamycin 50 μg ml^{-1}.

Single colonies were grown in 5 ml LB broth with appropriate antibiotic at 37°C with 200 rpm shaking. The A_{600} of overnight cultures was measured using a CE 2041 spectrophotometer set to 600 nm. Overnight cultures were diluted to an A_{600} of 0.05 in 10 ml fresh LB broth without antibiotics. Cultures were incubated at 37°C with 200 rpm shaking until they reached an A_{600} of 1.0. A 1 ml "whole cell" sample was taken. Cells were pelleted by centrifugation at 5,000 xg for 5 minutes. The supernatant was discarded and cells were resuspended in SDS-PAGE loading buffer with 140mM βME, standardised against culture A_{600} with a 1:75 dilution. This served as the late log phase sample.

The remaining 9 ml of culture was incubated overnight at 37°C with 200 rpm shaking. A 1 ml "whole cell" sample was taken and the A_{600} was measured. Cells were pelleted by centrifugation at 5,000 xg for 5 minutes. The supernatant was discarded and cells were resuspended in SDS-PAGE loading buffer with 140 mM βME, standardised against culture A_{600} with a 1:75 dilution. This served as the late stationary phase sample.

2.25 Small scale fractionation of cell extracts following osmotic shock

Small scale fractionation was carried out as previously described with slight modification (163, 235). Cells were incubated overnight at 37°C with 200 rpm shaking in 5 ml LB broth with appropriate antibiotics. The A_{600} of the overnight culture was measured and diluted to an A_{600} of 0.05 in 10 ml fresh LB broth without antibiotics.

Cells were incubated 37°C with 200 rpm shaking until they reach an A_{600} of 1.0. A 1 ml sample was taken and marked “whole cell”. A second 1 ml sample was taken and marked “insoluble”.

“Whole cell” fractions were centrifuged at 13,000 xg for 5 minutes to pellet the cells. Supernatant was collected and passed through a 0.2 µm Minisart hydrophilic cellulose acetate filter. Proteins in the supernatant were precipitated by the addition of 10% v/v TCA and 1% v/v Triton X-100 for 1 hour at 4°C. Precipitated protein was centrifuged at 17,000 xg, 4°C for 20 minutes. Protein pellets were resuspended in 1 ml ice cold 100% acetone (Sigma Aldrich) and vortexed vigorously to remove any remaining TCA. Precipitated protein was pelleted by centrifugation at 17,000 xg, 4°C for 20 minutes. Remaining acetone was allowed to evaporate. “Whole cell” and “supernatant” samples were resuspended in SDS-PAGE loading buffer with 140 mM βME. The “whole cell” fractions were standardised against culture A_{600} with a 1:75 dilution. These fractions served to provide the total protein in the cell. The “supernatant” fractions were standardised against culture A_{600} with a 1:50 dilution. These fractions served as a control to monitor different levels of protein export and secretion.

“Insoluble” fractions were centrifuged at 13,000 xg for 5 minutes to pellet the cells. The supernatant was discarded. Cells were gently resuspended in 250 µl of sterile resuspension buffer (0.5M sucrose, 40mM Tris-HCl, pH 7.4, 5mM EDTA). Lysozyme, chloride-free from chicken egg white (Sigma Aldrich) was added to a final concentration of 50 µg ml^{-1}. Sterile deionised water (230 µl) was added and cells were mixed by inverting the tube. The fractions were incubated for 5 minutes at room temperature. $MgCl_2$ (1 M, 10 µl) was added and the fractions were mixed gently. Cells were centrifuged for 10 minutes at 17,000 xg, 4°C. The supernatant was removed and protein was precipitated with 10% v/v TCA and 1% v/v Triton X-100 for 1 hour at 4°C. This was the “periplasmic” fraction.

Osmotically-shocked cells were resuspended gently in 500 µl of buffer (20 mM Tris-HCl, 5 mM EDTA, 1 tablet of cOmplete mini EDTA-free protease inhibitor cocktail (Roche)). DNaseI (1 U) was added and the tube was inverted gently on a rotary wheel (Fisher Brand) for 2 minutes or until viscosity decreased. $MgCl_2$ (1 M, 10 µl) was added

and the suspension was mixed by inversion. The fraction was centrifuged for 10 minutes at 17,000 xg, 4°C. The pellet was the “insoluble” fraction containing insoluble membrane and proteins. The supernatant was removed and protein was precipitated with 10% v/v TCA and 1% v/v Triton X-100 for 1 hour at 4°C. This supernatant was the “cytosolic-soluble” fraction.

The “periplasmic” and “cytosolic-soluble” fractions were centrifuged at 17,000 xg, 4°C for 20 minutes to pellet protein. Protein pellets were resuspended in 1 ml ice cold 100% acetone (Sigma Aldrich) and vortexed vigorously to remove any remaining TCA. Precipitated protein was pelleted by centrifugation at 17,000 xg, 4°C for 20 minutes. Remaining acetone was allowed to evaporate. The “periplasmic”, “cytosolic-soluble”, and “insoluble” fractions were resuspended in SDS-PAGE loading buffer with 140 mM βME, standardised against culture A_{600} with a 1:50 dilution.

2.26 FlhB$_{CC\ CT3xFLAG}$ secretion assay

S. enterica serovar Typhimurium SJW1103*flhB$_{CT3xFLAG}$* strains were inoculated into 10 ml LB broth for overnight growth. Overnight culture was diluted 1:100 into 100 ml LB broth and incubated at 37°C with 200 rpm shaking. When cells reached an A_{600} of 1.0, cells were pelleted by centrifugation at 10,000 xg for 10 minutes. The cell pellet was saved as the “whole cell” fraction. Supernatant was recovered and passed through a 0.2 μm Minisart hydrophilic cellulose acetate filter. Proteins in the supernatant were precipitated overnight at 4°C by the addition of 10% v/v TCA and 1% v/v Triton X-100. Precipitated protein was centrifuged at 17,000 xg, 4°C for 20 minutes. Protein pellets were resuspended in 5 ml ice cold 100% acetone (Sigma Aldrich) and vortexed vigorously to remove any remaining TCA. Precipitated protein was pelleted by centrifugation at 17,000 xg, 4°C for 20 minutes. “Whole cell” fractions were resuspended in 500 μl SDS loading buffer with 140 mM βME and “supernatant” fractions were resuspended in 250 μl SDS loading buffer with 140 mM βME.

2.27 Proline aminopeptidase activity assay

A previously described proline aminopeptidase activity assay was adapted for use with *Salmonella* cell lysate and protein (236). *E. coli* C41 (DE3) was transformed with pET15b and pET15b FliK$_{WT}$. Protein expression and HexaHis-tagged FliK purification

were as previously described. HexaHis-tagged $FliK_{WT}$ was dialysed overnight at 4°C against buffer G using 10 kDa cutoff regenerated cellulose dialysis membrane tubing (Spectrum Labs). C41 (DE3) cells with induced pET15b were resuspended in 1x PBS and lysed at 30 kpsi by cell disruption. Soluble fraction was collected by centrifugation at 40,000 xg, 4°C for 1 hour. The soluble fraction from induced C41 (DE3) pET15b served as the positive peptidase control for this experiment as *E. coli* harbors native proline aminopeptidases.

L-proline p-nitroanilide trifluoroacetate salt (Sigma Aldrich) was resuspended to 2 mM in either PBS or buffer G. C41 (DE3) pET15b soluble lysate (500 μl) with or without overexpressed HexaHis-tagged $FliK_{WT}$ (500 μl) was mixed with 500 μl of 2 mM L-proline p-nitroanilide trifluoroacetate salt. 500 μl of PBS or buffer G mixed with 500 μl of 2mM L-proline p-nitroanilide trifluoroacetate salt served as negative peptidase controls. Samples were incubated for 30 minutes at either 20°C or 50°C. The reactions were stopped with 1 ml 30% v/v acetic acid. Sample (1 ml) was added to a 10x4x45 mm polystyrene cuvette and analysed using a Cecil CE 2041 spectrophotometer set to 405 nm.

2.28 In vitro $FlhB_{CC}$ dissociation assay

$FlhB_{C\ Ct3xFLAG}$ fused to inteinCBD was purified as described in 2.12.2, but dialysed into buffer B. $FlhB_{C\ Ct3xFLAG}$ was diluted in buffer B to 10 μM and frozen in liquid nitrogen for storage at -80°C.

E. coli C41 (DE3) was transformed with pET15b or pET15b-FliK (HexaHis-FliK). Proteins were expressed as described in 2.11.1. Cell pellets from both were resuspended in 25 ml buffer B supplemented with 2 U DNaseI and 1 tablet of cOmplete protease inhibitor cocktail, EDTA-free. Cells were lysed at 30 kpsi using cell disruption. Soluble lysate was collected by centrifugation at 60,000 xg, 4°C for 1 hour.

The soluble lysates were transferred to 50 ml falcon tubes (Greiner). $FlhB_{C\ Ct3xFLAG}$ was diluted to 1 μM. $FlhB_{C\ Ct3xFLAG}$ and soluble lysates were incubated at 20°C on a rotary wheel. Samples (1 ml) were taken at 0, 30, 90, and 180 minutes. Samples were precipitated for 1 hour at 4°C with the addition of 10% v/v TCA and 1% v/v Triton-

X100. Precipitated protein was resuspended by vigorous vortexing in 100 μl SDS loading buffer supplemented with 140 mM βME. Samples were analysed by immunoblotting. Mouse PentaHis antibody (Qiagen) was used to detect FliK (HexaHis-FliK) and mouse anti-FLAG M2 antibody (Sigma) was used to detect $FlhB_{C\ Ct3xFLAG}$.

Chapter 3

Early subunit interactions with the export gate component, FlhB

3.1 Introduction

Unlike the late flagellar subunits, early subunits for the rod and hook do not possess chaperones to traffic them to the flagellar export machinery (126). Recent work has shown that early subunits bind via their gate recognition motif (GRM) to the cytoplasmic domain of the export gate component, FlhB (86).

The N-terminal GRM is highly conserved among early subunits and across bacterial species possessing a flagellum (86). The motif starts with a phenylalanine (F), a residue found across early subunits and bacterial species. This phenylalanine is followed by three residues of varying composition and a large hydrophobic residue in the fifth position (86). It is predicted that the phenylalanine and the large hydrophobic residue in the fifth position putatively interact with the surface-exposed hydrophobic residues on $FlhB_{CC}$. The function of the three middle residues is unknown, but may serve to place the GRM properly in the binding pocket.

Evans and colleagues identified four hydrophobic residues on $FlhB_C$ that are important for subunit binding *in vitro*. These residues – two alanines, a leucine, and a proline – were individually mutated into glutamates, replacing hydrophobic residues with a large charged residue to alter the binding site to prevent early subunit binding. Any individual replacement was enough to abolish subunit interaction with FlhB as observed by a protein pulldown assay (86). All four residues are located on $FlhB_{CC}$. It isn't clear how the surface-exposed hydrophobic residues on $FlhB_{CC}$ interact directly with subunit GRM or if they may play some other role in subunit export.

Although we are not certain how tightly early subunits bind to FlhB, we know the interaction is relatively weak. Isothermal titration calorimetry data for one such early subunit, FlgD (the hook cap monomer), shows that affinity between the hook cap and $FlhB_C$ is only around 40 μM. This weak interaction between subunit and gate allows the subunit to be removed from $FlhB_C$ to be exported and transit to the distal site of assembly in the nascent flagellum (86). Morris and colleagues used surface plasmon

resonance to observe the interaction between full length FliK and autocleaved and non-autocleaved variants of $FlhB_C$. They also reported micromolar affinities, although on the order of 1-10 μM, with very little change in affinity or reaction kinetics between autocleaved and non-autocleaved variants of $FlhB_C$ (188, 191). This interaction between the flagellar molecular ruler and early subunit export gate is much stronger than the reported affinity (430 μM) of the *Yersinia* injectisome molecular ruler to the early subunit export gate (89).

The hook cap must first assemble to coordinate hook assembly, but, as the hook is a relatively short structure, presumably only one capping structure is needed to control hook assembly. The molecular ruler, however, must measure the hook throughout assembly and keep the length within a narrow range if the flagellum is to function properly. The difference between affinities of hook cap and the molecular ruler for the flagellar early subunit export gate is approximately five-fold, but it is unknown why these two subunits, with such similar GRMs, bind differently to FlhB (86, 191). Our hypothesis is that these affinities may play a role in establishing export order and stoichiometry of exported subunit.

The work in this chapter seeks to further analyse the role of early subunit GRMs for efficient subunit export. Early subunit binding affinities to $FlhB_C$ were evaluated *in vitro* to test whether these affinities may impose order on subunit export. Finally, crystallography and nuclear magnetic resonance (NMR) spectroscopy were used to characterise the binding pocket environment on $FlhB_{CC}$ upon early subunit binding and investigate any further structural changes at $FlhB_C$ caused by early subunit binding, but not detected by the published pulldown assay results.

3.2 Mutations of the gate recognition motif impair early subunit export

It is known that deleting the GRM from early subunits abolishes the *in vitro* interaction between early subunits and $FlhB_C$ and attenuates subunit export *in vivo*. Presumably, this is because subunit is no longer able to bind to the export gate and therefore cannot be exported efficiently, affecting the cell's ability to swim (86). It is not known, however, which of the GRM residues are most important for subunit export and swimming motility.

The hook subunit, FlgE, is unusual among *Salmonella* early subunits in that its GRM contains two phenylalanine residues (Figure 3.1A). Recombinant pBAD18amp plasmids carrying mutations of the GRM of *flgE* under the control of the P_{BAD} promoter were constructed for expression in a *flgE* null strain of *Salmonella* Typhimurium to measure the mutants' ability to swim and export hook subunit. Mutations were either conservative, replacing the first phenylalanine and fifth phenylalanine in the FlgE GRM with an alanine, or non-conservative, replacing these GRM residues with lysine (Figure 3.1A). All recombinant subunits contained a three-repeating FLAG tag (3xFLAG) sequence between residues 234 and 235 to allow for better detection of subunit export.

A *flgE* null strain of *Salmonella* Typhimurium was transformed with the recombinant plasmids to measure the effect of GRM mutation on swimming motility. Some of the mutated GRMs could complement the loss of chromosomal *flgE in trans* (Figure 3.1B). Cells expressing the FlgE $F_{43}K$ and $F_{43}A$ variants were as motile as wild type cells, whereas cells expressing the $F_{39}K$, $F_{39}A$, and $F_{39}A$, $F_{43}A$ (FFAA) variants of FlgE were only weakly motile. Cells expressing the FlgE ΔGRM and FlgE $F_{39}K$, $F_{43}K$ (FFKK) variants were not motile. None of the mutations were dominant when expressed in wild type *Salmonella* SJW1103 (Figure 3.1C).

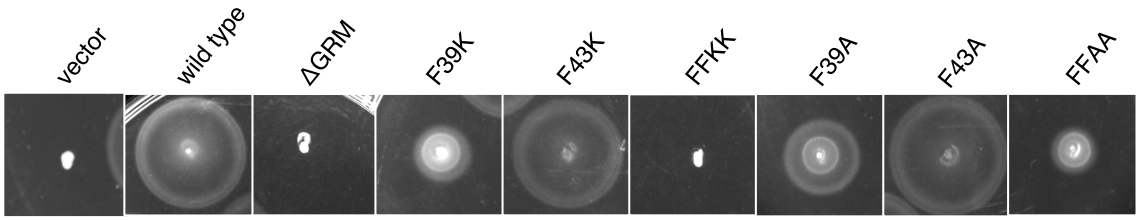
Why do only some of the GRM mutations complement swimming motility? We hypothesised that the effect on swimming motility was due to inefficient subunit export. To measure this, a *flgD* null strain of *Salmonella* Typhimurium was transformed with the recombinant plasmids described above. This strain would not assemble hook monomers into a hook structure, but would still be capable of exporting them into the culture supernatant. FlgE wild type and the GRM mutants expressed similarly from the plasmid after induction with 0.02% w/v L-arabinose, confirming that the mutations did not affect protein expression (Figure 3.2).

Only the FlgE $F_{43}A$ variant approached levels of export comparable to wild type cells, nearly 80% of wild type FlgE export, although that difference was not statistically significant (p-value 0.07). The $F_{39}K$, $F_{43}K$, $F_{39}A$, and FFAA variants exhibited similarly impaired export, only around 40-50% of wild type export. In comparison, the GRM

A

	FlgE GRM
Wild type	$_{39}$FADMF$_{43}$
F39K	$_{39}$KADMF$_{43}$
F43K	$_{39}$FADMK$_{43}$
FFKK	$_{39}$KADMK$_{43}$
F39A	$_{39}$AADMF$_{43}$
F43A	$_{39}$FADMA$_{43}$
FFAA	$_{39}$AADMA$_{43}$

B Complementation of SJW1103*ΔflgE::FRT*



C Dominance in SJW1103

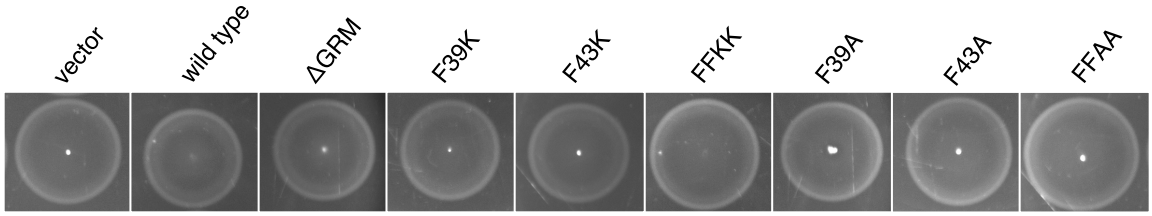


Figure 3.1 Replacing the first and fifth residue of the GRM attenuates cell motility

(A) FlgE GRM mutants used in this study.

(B) Complementation of SJW1103*ΔflgE* transformed with pBAD18amp plasmids carrying *flgE*$_{mid3xFLAG}$ wild type and GRM variants on swimming motility agar (0.25% w/v agar) with 0.02% w/v L-arabinose after incubation for four hours at 37°C. No motility was observed in the absence of arabinose. Vector is empty pBAD18amp.

(C) Dominance of SJW1103 transformed with pBAD18amp plasmids carrying *flgE*$_{mid3xFLAG}$ wild type and GRM variants on swimming motility agar (0.25% w/v agar) with 0.02% w/v L-arabinose after incubation for four hours at 37°C. Vector is empty pBAD18amp.

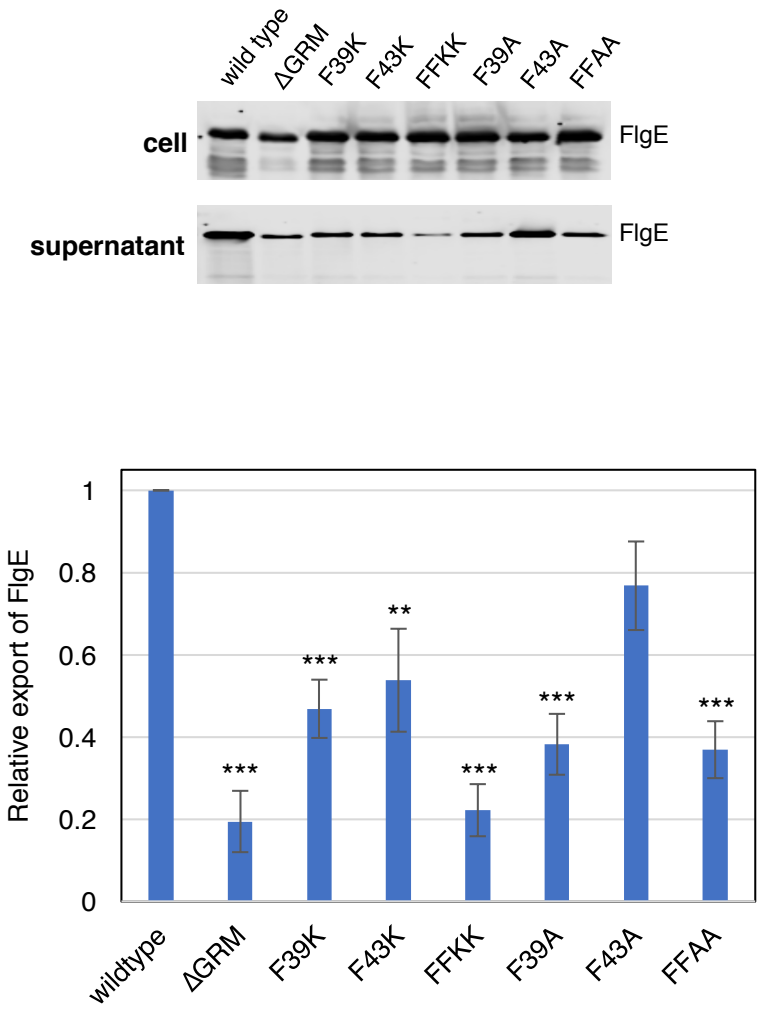


Figure 3.2 Replacing the first and fifth residue of the GRM attenuates subunit export

Export data of *flgE*$_{mid3xFLAG}$ wild type and GRM variants carried on pBAD18amp. Plasmids were transformed into SJW1103*ΔflgD* and protein expression was induced by the addition of 0.02% w/v L-Arabinose. Whole cell fractions were collected by centrifugation and diluted 1:100 in SDS loading buffer based on culture density as measured by absorbance at 600 nm. Supernatant fractions were collected by centrifugation, filtered to remove any cells, and TCA precipitated to concentrate secreted protein. Samples were diluted 1:50 in SDS loading buffer based on culture density as measured by absorbance at 600 nm. Mouse anti-FLAG M2 (Sigma) anitbody was used to detect FlgE in both fractions. Secondary infrared antibody, Donkey anti-mouse (Li-Cor), was used to detect the anti-FLAG antibody.

The relative levels of export for each variant compared to wild type FlgE are shown. The error bars represent the standard error of the mean, which was calculated in Excel from four biological replicates and analysed using a Li-Cor Odyssey CLx and Image Studio Lite (Li-Cor). Wild type FlgE export was counted as 1 in each replicate. ** indicates a p-value < 0.01. *** indicates a p-value < 0.001.

deletion (ΔGRM) and the FFKK variant led to a remarkable decrease in export compared to wild type, around 20% of wild type export. The similarity between the full deletion of the GRM and the non-conservative replacement of the first and fifth residues suggests that it is these two residues which are most important for subunit export.

Statistical analysis of the export data confirms that the reduction of subunit export due to the FlgE GRM mutations, except FlgE F43A, compared to wild type FlgE export is significant with p-values below 0.01. To determine this, export of the GRM mutants was compared to FlgE wild type, which was standardised to a relative export level of 1.0 across all four biological replicates. The standard error of the mean was calculated for each sample and is represented by the error bars shown (Figure 3.2). An F test was performed to determine variance across samples before a Student's t-test was performed.

The *in vivo* results in this section expand on the previously reported findings that the first and fifth residues of the early subunit GRM are vital for subunit export and swimming motility. Next, I investigated the strength of the interaction between the FlgE GRM and $FlhB_C$, incorporating the point mutants used *in vivo* to determine how they effected subunit binding affinity.

3.3 The hook protein, FlgE, binds weakly to $FlhB_C$

To study the *in vitro* interactions between the FlgE GRM and $FlhB_C$, $FlhB_C$ was first purified as described previously with slight modifications (Figure 3.3A and 3.3B) (223). After overnight elution from chitin resin, $FlhB_C$ was further purified by cation exchange chromatography using 50 mM Tris-HCl, pH 7.4, buffer and a gradient of NaCl (up to 1.2M) to elute $FlhB_C$ from the column.

Concentrated fractions from cation exchange chromatography were pooled and applied to a S75 10/300 superdex gel filtration column (GE Healthcare), pre-equilibrated into the appropriate buffer. $FlhB_C$ eluted in two peaks (Figure 3.3C). The first peak contained both $FlhB_C$ and contaminating bands and was discarded. The second, large peak contained only wild type, cleaved $FlhB_C$ (Figure 3.3C), which was

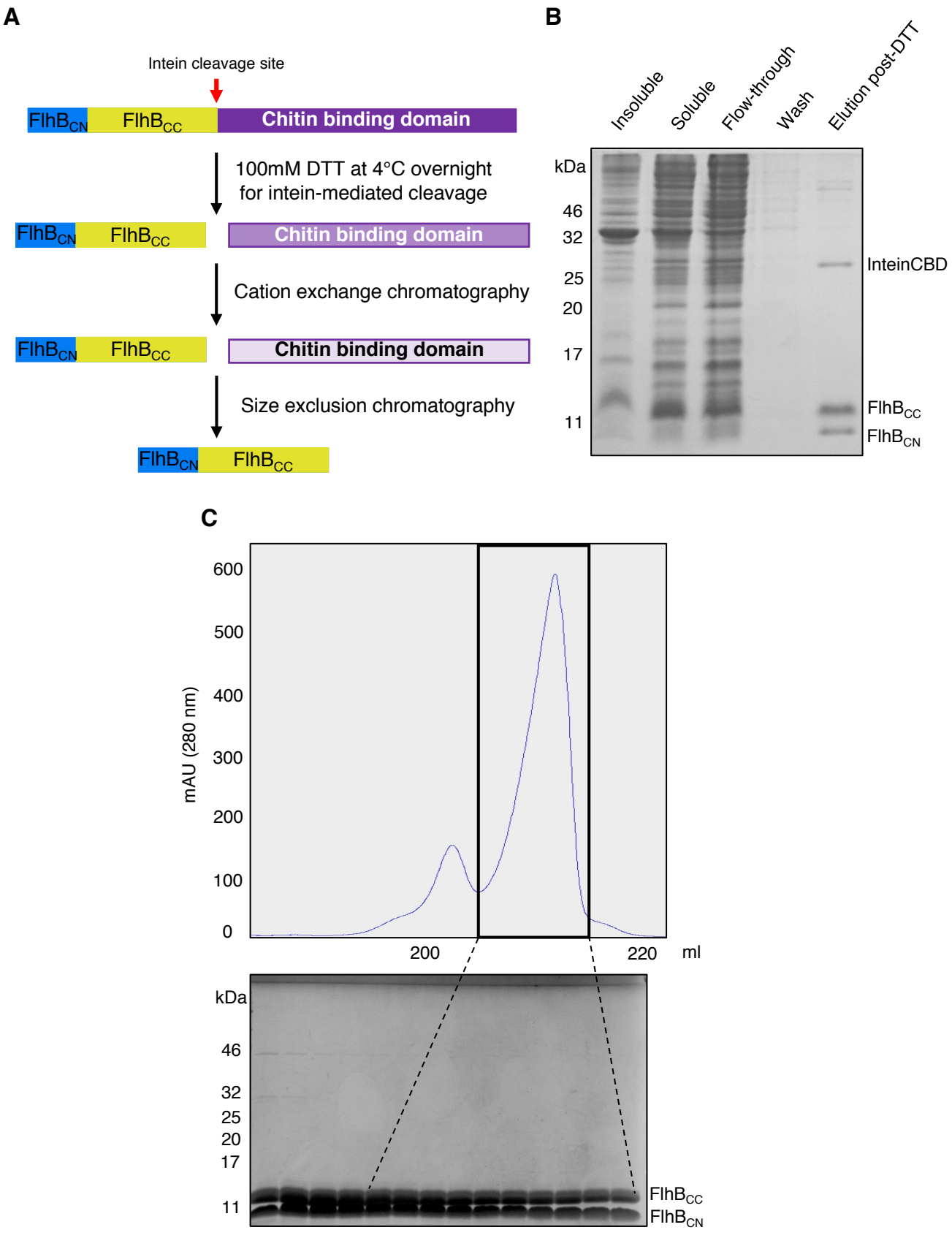


Figure 3.3 Purification procedure of FlhB$_C$ residues 219-383, G$_{383}$A using inteinCBD

(A) FlhB$_C$ is fused to an intein-chitin binding domain (inteinCBD) tag and initially purified on chitin resin (NEB). FlhB$_C$ is cleaved from inteinCBD and eluted using Buffer A supplemented with 0.1 M Dithiothreitol (DTT) after overnight incubation at 4°C.

(B) FlhB$_C$ after elution from chitin column. Most of the eluate was FlhB$_C$, but there was contamination with intein-CBD. FlhB$_C$ was further purified using cation exchange chromatography and size exclusion chromatography.

(C) Fractions from the second elution peak were pooled and concentrated to desired molarity.

collected and pooled before being concentrated to appropriate molarity for subsequent experiments. As previously reported, the $FlhB_{CC}$ domain remains non-covalently associated with the $FlhB_{CN}$ domain in solution (190, 192).

Synthetic peptides were then designed to mimic FlgE GRM binding to $FlhB_C$ *in vitro* (Covalab S.A.S, France) (Figure 3.4A). Each peptide contains the wild type or variant GRMs described above in the middle of the peptide, flanked by the native residues found in *Salmonella* Typhimurium FlgE. Peptides easily went into solution in deionised water (MilliQ) and stocks were frozen in liquid nitrogen and stored at -20°C.

Isothermal titration calorimetry (ITC) was performed for peptides containing the wild type or variant FlgE GRM peptides to assess the affinity for hook subunit GRM to $FlhB_C$. Peptide was applied in a ten-fold molar excess to $FlhB_C$ by direct injection. Wild type FlgE GRM peptide bound to $FlhB_C$ with a K_d of around 10 μM (Figure 3.4B). This is a stronger interaction than that reported for full length FlgD binding to $FlhB_C$, which has a K_d of 39 μM (86). The peptides containing variant FlgE GRMs, with either conservative or non-conservative replacements of either the first or fifth GRM residues, were unable to bind FlhB in a range measurable by ITC, corroborating the *in vivo* data previously presented that showed these mutations impaired hook subunit export (Figure 3.4C-F).

3.4 Early subunits bind transiently to $FlhB_C$

The FlgE GRM peptides bound $FlhB_C$ with a K_d approximately four-fold lower than full-length FlgD. Additional peptides were designed for all the early subunits to measure the K_d upon binding to $FlhB_C$ using ITC. We hypothesized that the K_d between early subunits and $FlhB_C$ determines subunit export order (Figure 3.5A and 3.6A). All peptides save for the FlgC and FlgD GRM peptides were easily resuspended in deionised water (MilliQ). FlgD GRM peptide later went into solution when diluted into a Tris-salt buffer, whilst FlgC GRM peptide required the addition of 10% v/v dimethyl sulfoxide (DMSO, Sigma).

The early subunit GRM peptides bound to $FlhB_C$ with micromolar affinities, ranging from approximately 5 to 45 μM. The hook subunit GRM peptide, FlgE, and the

A

Peptide	Sequence
FlgE	KSGTASFADMFAGSKVGLG
FlgE $F_{39}K$	KSGTASKADMFAGSKVGLG
FlgE $F_{39}A$	KSGTASAADMFAGSKVGLG
FlgE $F_{43}K$	KSGTASFADMKAGSKVGLG
FlgE $F_{43}A$	KSGTASFADMAAGSKVGLG

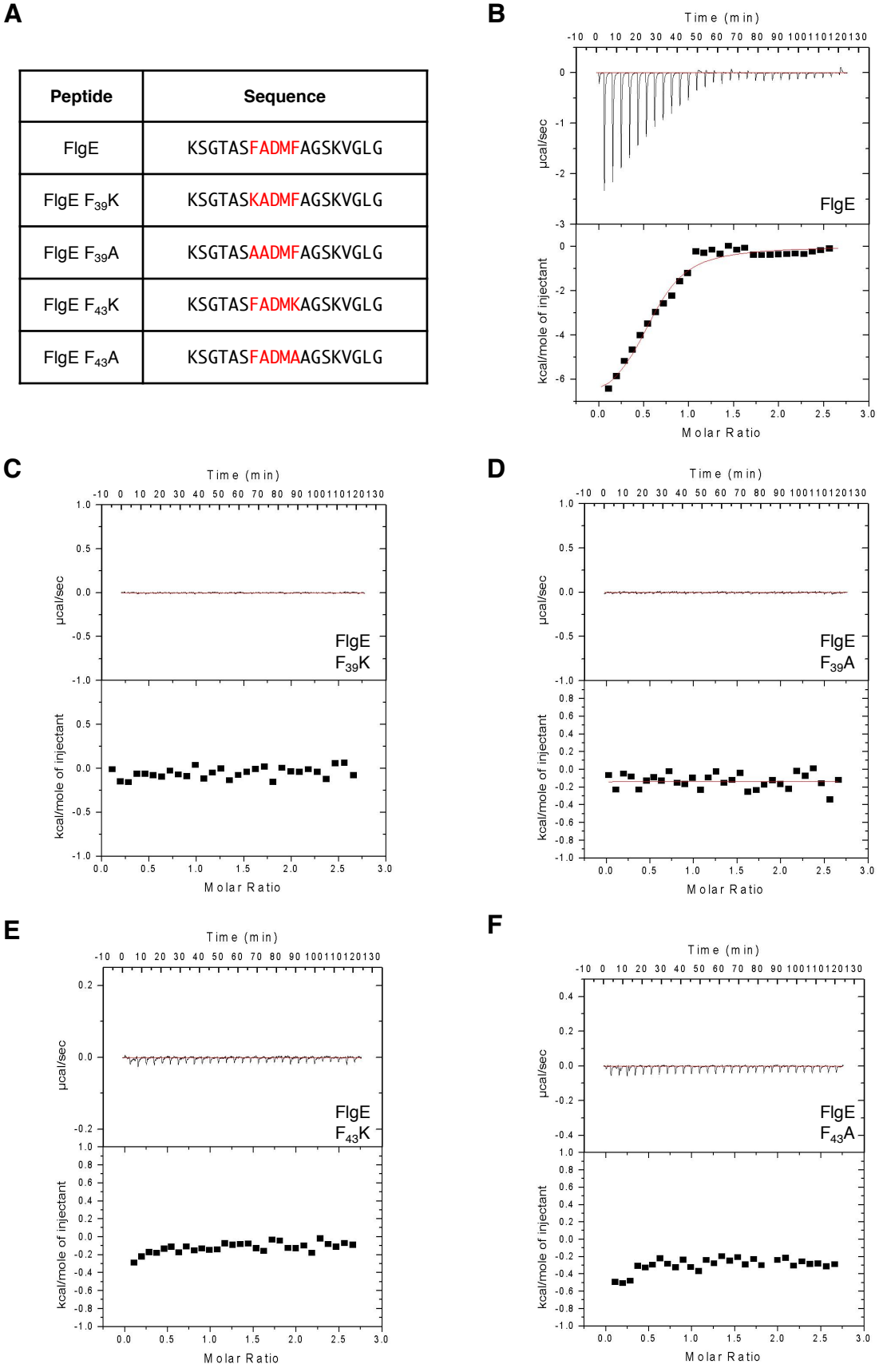


Figure 3.4 Single residue changes in a FlgE GRM peptide abolish binding to $FlhB_C$

(A) Table of synthetic peptides used in the ITC studies with $FlhB_C$. Both protein and peptide were kept in degassed 50 mM Tris, pH 7.4 150 mM NaCl buffer at 25°C in preparation for ITC.

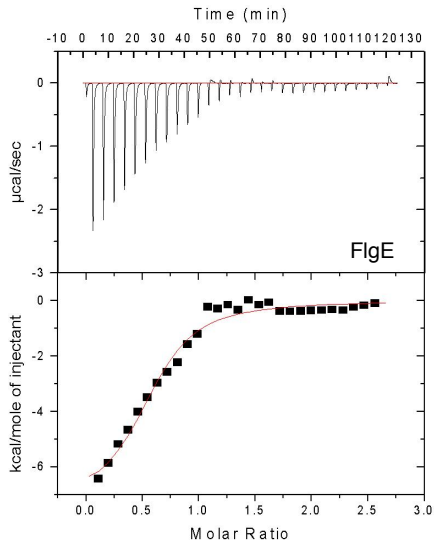
(B) ITC trace of $FlhB_C$ titrated with FlgE wild type. The heat of dilution of FlgE wild type peptide in buffer was subtracted from $FlhB_C$ titrated with FlgE wild type to generate the best fit line.

(C-F) ITC traces of $FlhB_C$ titrated with FlgE GRM mutant peptides.

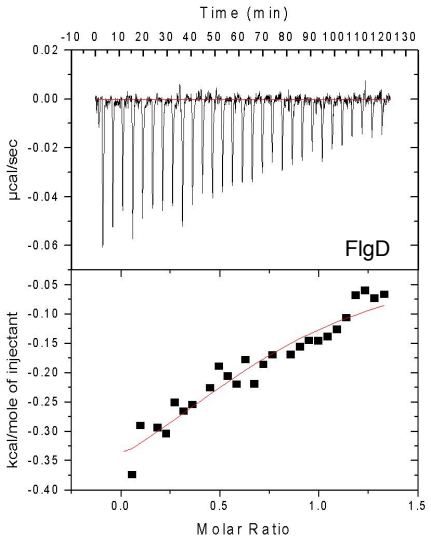
A

Peptide	Sequence
FlgE	KSGTASFADMFAGSKVGLG
FlgD	LQSSFLTLLVAQLKNQD
FliK	KTTGSAEDFLALLAGALGAD

B



C



D

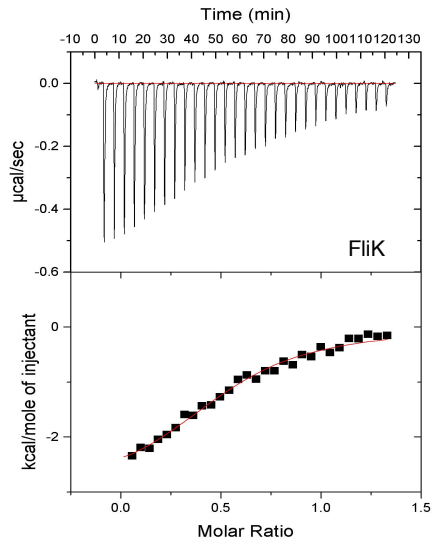


Figure 3.5 Hook subunits bind with varying affinities to $FlhB_C$

(A) Table of synthetic flagellar rod peptides used in the ITC studies. Both protein and peptide were kept in degassed 50 mM Tris, pH 7.4 150 mM NaCl buffer at 25°C in preparation for ITC.

(B-D) ITC traces of $FlhB_C$ titrated with hook subunit GRM peptides. The heat of dilutions for peptide in buffer was subtracted from $FlhB_C$ titrated with each peptide to generate the best fit line. The GRM peptides used in each experiment are listed on each trace.

A

Peptide	Sequence
FliE	THSQSTVSFAGQLHAALDR
FlgB	YQARDIDFASELKKVMVRG
FlgC	YRAKQVVFQVDAAPGQATGG
FlgF	NASTPGFRAQLNALRAV
FlgG	KRQRAVFEDLLYQTIRQ

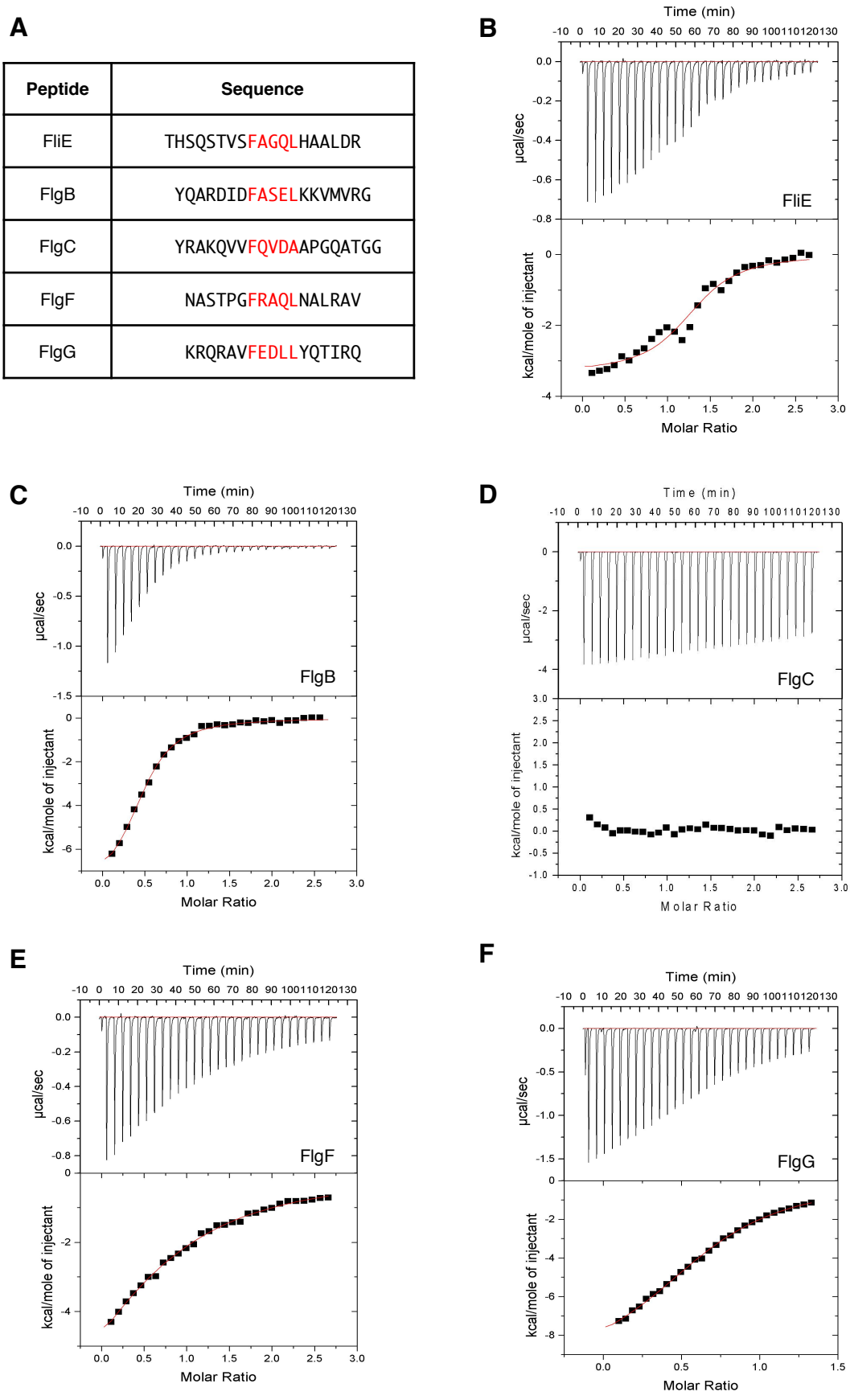


Figure 3.6 Rod subunits bind with varying affinities to $FlhB_C$

(A) Table of synthetic flagellar rod peptides used in the ITC studies. Both protein and peptide were kept in degassed 50 mM Tris, pH 7.4 150 mM NaCl buffer at 25°C in preparation for ITC.

(B-F) ITC traces of $FlhB_C$ titrated with rod subunit GRM peptides. The heat of dilutions for peptide in buffer was subtracted from $FlhB_C$ titrated with each peptide to generate the best fit line. The GRM peptides used in each experiment are listed on each trace. FlgC GRM peptide was measured in the presence of 10% v/v DMSO.

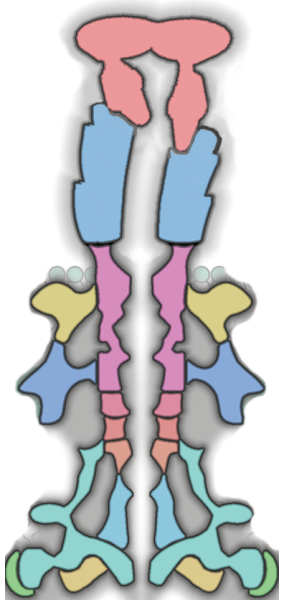
secreted molecular ruler GRM peptide, FliK, have K_d's of intermediate strength, between 10 and 15 μM (Figure 3.5B and 3.5D). The FlgD peptide, the hook cap, binds with a K_d of 38.8 μM, in line with the reported K_d for full-length FlgD (Figure 3.5C). The strongest binders were those of the proximal rod subunit GRM peptides, FliE and FlgB, which both share a similar affinity for $FlhB_C$ (Figure 3.6B and 3.6C). The distal rod subunit GRM peptides, FlgF and FlgG, are weaker binders by approximately eight- and five-fold, respectively (Figure 3.6E and F). All told, there is roughly a ten-fold difference across the K_d's of all early subunit GRM peptides (Figure 3.7).

3.5 How might FlgC be exported

Notably, the GRM peptide for the FlgC proximal rod subunit did not bind to $FlhB_C$ in a magnitude measurable by ITC. This peptide was relatively hydrophobic and precipitated in the ITC experimental buffer. For this reason, buffer supplemented with 10% v/v DMSO was used to keep the FlgC GRM peptide in solution, but still no binding to $FlhB_C$ could be observed in that experiment (Figure 3.8A). FlgB GRM peptide in buffer with 10% v/v DMSO was also titrated against $FlhB_C$ and the heat of binding was measured by ITC as a control to measure any effect of DMSO on peptide-$FlhB_C$ binding. The K_d was altered slightly in this experiment, from 5.6 μM to 9.6 μM, but binding was still observed (Figure 3.8B). That data indicate that DMSO does not substantially weaken subunit binding to $FlhB_C$.

Is the FlgC GRM not important for swimming motility? First, a sequence alignment of FlgC from seven species was performed to compare putative FlgC GRM sequences (Figure 3.9A). The closely related *E. coli*, *S. typhimurium*, and *Y. enterocolitica* all showed similar GRM sequences. As expected, the first phenylalanine of the GRM was also highly conserved across all species in the alignment. The phenylalanine of FlgC GRM of *B. subtilis* appears later in the protein sequence, at residue 60, compared to other species and is followed by a leucine in the fifth position. *A. fabrum* also has a leucine in the fifth position of the FlgC GRM. *B. burgdorferi* and *A. aeolicus* shared a valine in the fifth position (Figure 3.9A).

FlgC GRMs vary slightly across species, but the enteric organisms all share an alanine in the fifth position. When the fifth residue of the FlgE GRM was mutated into alanine

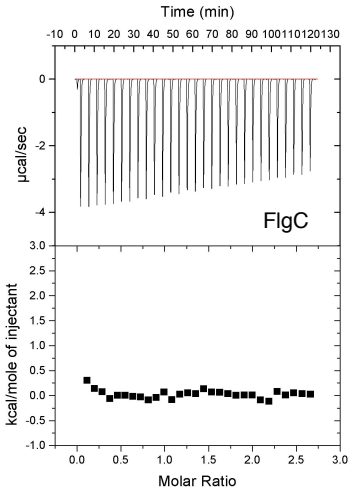


	K_d μM
FlgD	38.80
FlgE	10.80
FliK	15.96
FlgG	31.80
FlgF	48.56
FlgC	not determined
FlgB	5.60
FliE	5.72

Figure 3.7 Rod and hook subunits bind transiently to $FlhB_C$

Summary of dissociation constants of GRM peptides binding to $FlhB_C$ using ITC. Subunits are listed in order of final assembly, starting from FliE up to FlgD, and colour-coded to refer to the rod and hook structure. FliK is an exported subunit that is not assembled into the final structure.

A



B

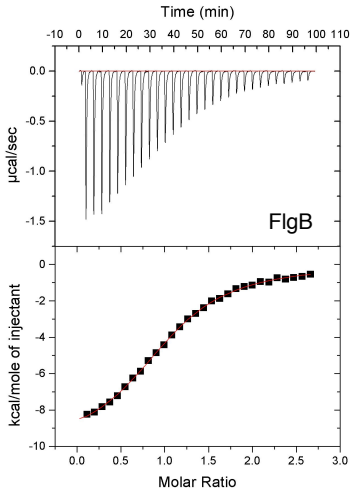


Figure 3.8 DMSO does not substantially affect the $FlhB_C$-GRM interaction as measured by ITC

(A) ITC trace of $FlhB_C$ titrated with FlgC GRM peptide in the presence of 10% v/v DMSO. The heat of dilutions for peptide in buffer was subtracted from $FlhB_C$ titrated with FlgC to generate the best fit line. The K_d could not be determined from this data, suggesting that ITC cannot measure the interaction between the FlgC peptide and $FlhB_C$.

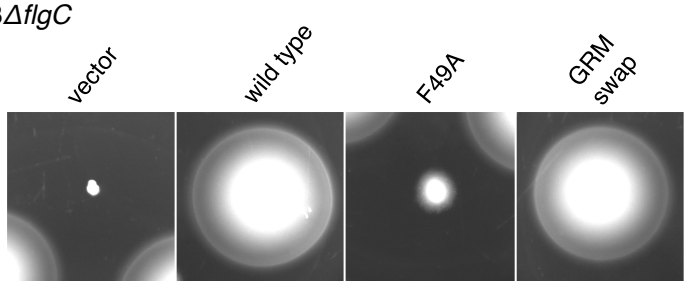
(B) ITC trace of $FlhB_C$ titrated with FlgB GRM peptide in the presence of 10% v/v DMSO. The heat of dilutions for peptide in buffer was subtracted from $FlhB_C$ titrated with FlgB to generate the best fit line. The K_d was recorded as 9.6 μM in the presence of DMSO. The K_d for the FlgB GRM peptide-$FlhB_C$ interaction without DMSO is 5.6 μM.

A

FlgC GRM

Aquifex	$_{38}$DGTPFNRLVPVFEAVENLEDPSE
Salmonella	$_{45}$KQVVFQVDAAPGQATGGVKVASV
E. coli	$_{45}$KQVVFQVNAAPGAATGGVKVADV
Yersinia	$_{45}$KQVVFQVAAQPGQETGGVRVAQI
Borrelia	$_{45}$QRIIFAPRVNNPYWKGPFIPDYL
Agrobacterium	$_{47}$KTVTFGSELDHVSGVERVKVKKL
Bacillus	$_{56}$KGESFSSILNSQMSGSGNAGNGV

B Complementation of SJW1103*ΔflgC*



C Dominance in SJW1103

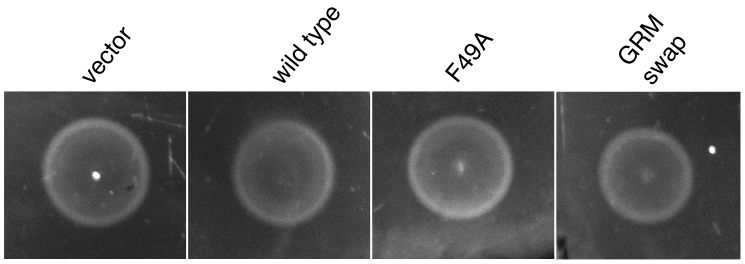


Figure 3.9 The conserved phenylalanine of the FlgC GRM is important for cell motility

(A) Sequence alignment of the proximal rod protein FlgC from selected species with the proposed GRM coloured red. Possible second GRM sites are coloured blue. Sequences were acquired from UniProt and aligned with Clustal Omega using default settings. *Aquifex aeolicus* strain VF5; *Salmonella enterica* serovar Typhimurium strain LT2; *Escherichia coli* strain K12; *Yersinia enterocolitica*; *Borrelia burgdorferi* strain ATCC 35210; *Agrobacterium fabrum* strain C58; *Bacillus subtilis* strain 168.

(B) Complementation of SJW1103*ΔflgC* transformed with pBAD18amp plasmids carrying *flgC* wild type and GRM mutants on swimming motility agar (0.25% agar) with 0.002% L-arabinose after incubation for 18 hours at 30°C. Vector is empty pBAD18amp. GRM swap is FlgC carrying the GRM sequence from FlgB (FASEL) rather than its native sequence (FQVDA).

(C) Dominance of SJW1103 transformed with pBAD18amp plasmids carrying *flgC* wild type and GRM mutants on swimming motility agar (0.25% agar) with 0.2% L-arabinose after incubation for four hours at 37°C. Vector is empty pBAD18amp.

in the previous experiments, cells expressing this variant could still swim and FlgE subunit export was only minimally affected. This mutation was unable to bind $FlhB_C$, however, suggesting that FlgC may be exported through another mechanism.

Recombinant pBAD18amp plasmids encoding FlgC GRM variants under the control of the P_{BAD} promoter were constructed to test whether the FlgC GRM was necessary for motility. Cells expressing wild type *Salmonella flgC* off the plasmid could complement the loss of *flgC* from the chromosome, but cells expressing an alanine replacement of the conserved phenylalanine in the FlgC GRM ($F_{49}A$) were not as motile (Figure 3.9B), the cells only weakly swimming. Interestingly, cells expressing a *flgC* variant where the *flgC* GRM was replaced with the GRM from *flgB* (FQVDA→FASEL), was as motile as wild type FlgC (Figure 3.9B). Recombinant plasmids with these *flgC* variants were also introduced into wild type *Salmonella* to test whether they were dominant over wild type *flgC*. No dominance effect could be seen (Figure 3.9C).

I hypothesised that FlgC may natively be co-exported with FlgB as prior results have suggested an interaction between the two subunits (132, 133). Recombinant pTXBI plasmid was constructed to co-express HexaHis-3xFLAG-tagged FlgB and FlgC, which was directly fused to the intein-chitin binding domain found on the plasmid. FlgB expression was under the control of the vector-encoded T7 promoter, while FlgC was expressed under its native promoter found immediately upstream of *flgC* at the end of *flgB*. This allowed for high expression, but also maintained the ratio of FlgB:FlgC expression found in the cell.

Three expression conditions were tested (Figure 3.10A). In all cases, protein was induced by the additional of 1 mM IPTG to the culture as described in Materials and Methods. HexaHis-3xFLAG-FlgB expressed well under all conditions, with overnight growth at 20°C optimal. FlgC-inteinCBD, however, was only produced when expression cultures were grown overnight at 20°C (Figure 3.10A). This latter condition was selected for future protein expression.

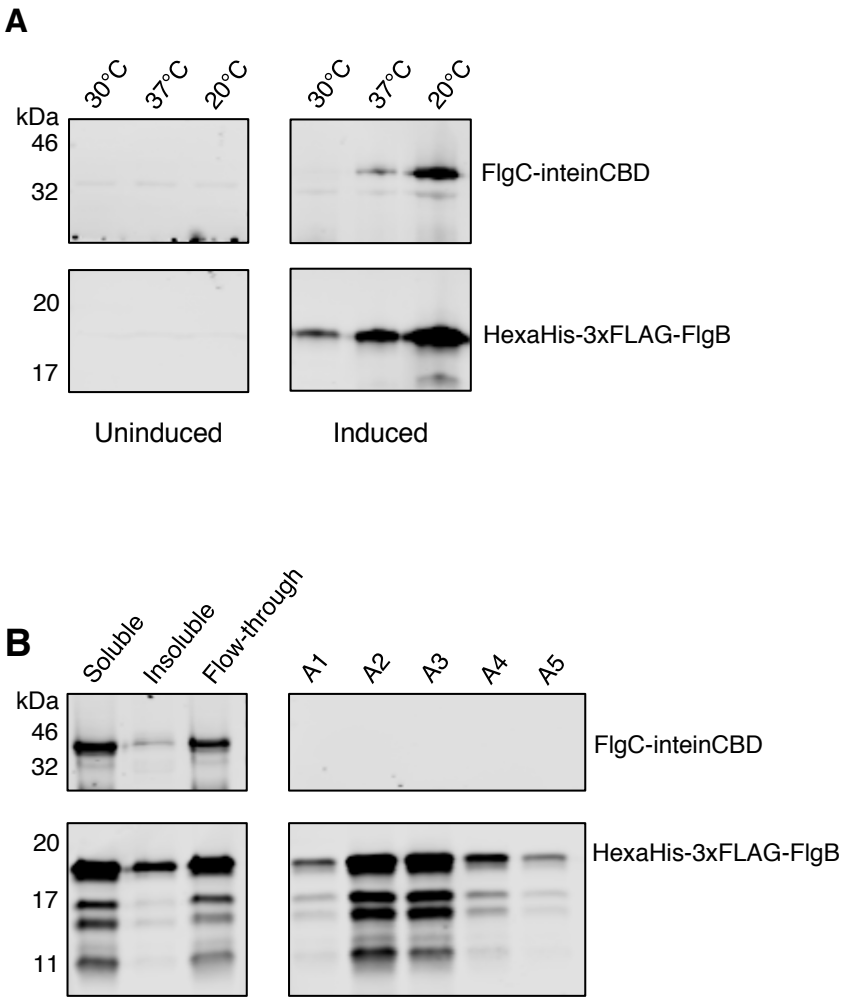


Figure 3.10 FlgB and FlgC were co-expressed but could not be co-purified

(A) Overnight, low temperature incubation was optimal for FlgB and FlgC co-expression. Co-expression of HexaHis-3xFLAG-FlgB and FlgC-inteinCBD from the pTXBI vector (NEB). Three conditions were tested, all with 2YT medium and induction with 1 mM IPTG. Cells grown at 30°C and 37°C with 200 rpm shaking were induced at A_{600} 1.0 and protein expression was carried out for 4 hours. Cells grown at 20°C with 200 rpm shaking were induced at A_{600} 1.0 and protein expression was carried out for 20 hours. Whole cell fractions (1 ml) were collected by centrifugation and diluted 1:500 in SDS loading buffer based on culture density as measured by absorbance at 600nm. Mouse anti-FLAG M2 (Sigma) antibody was used to detect FlgB and mouse anti-CBD (NEB) was used to detect FlgC. Secondary infrared antibody, donkey anti-mouse (Li-Cor), was used to detect the primary antibodies. Samples were imaged using a Li-Cor Odyssey CLx and Image Studio Lite (Li-Cor).

(B) FlgB and FlgC could not be purified from a nickel column. The cell pellet containing expressed recombinant protein was resuspended in 20 mM Tris-HCl, pH 8.0 buffer with 20 mM imidazole and 150 mM NaCl. Cells were lysed using a cell disruptor at 30k psi and the soluble fraction was collected by centrifugation. Co-purification of HexaHis-3xFLAG-FlgB and FlgC-inteinCBD was attempted on a 1 ml HisTrap Excel column (GE Healthcare) attached to a BioLogic FPLC (Bio-Rad). Soluble protein was loaded using 0.5 ml min^{-1} flow rate. Elution fractions A1-A5 (1 ml each) were collected by eluting bound protein with 300 mM imidazole. Samples (0.1 ml) were diluted 1:2 in SDS loading buffer and analysed by immunoblotting. Mouse anti-FLAG M2 (Sigma) antibody was used to detect FlgB and mouse anti-CBD (NEB) was used to detect FlgC. Secondary infrared antibody, donkey anti-mouse (Li-Cor), was used to detect the primary antibodies. Samples were imaged using a Li-Cor Odyssey CLx and Image Studio Lite (Li-Cor).

Both proteins were soluble in the purification buffer. To test whether FlgB and FlgC may interact *in vitro*, I attempted to co-purify the proteins using a nickel column attached to an FPLC. A low flow rate was used to minimise column pressure and not break any interaction between the two proteins. Only FlgB was recovered from fractions eluted off the nickel column using buffer supplemented with 300 mM imidazole. FlgC did not bind the column or FlgB and was only found in the flow-through fraction (Figure 3.10B). This data suggests that FlgC may not interact with FlgB, or does so too weakly to be observed by co-purification using the described method.

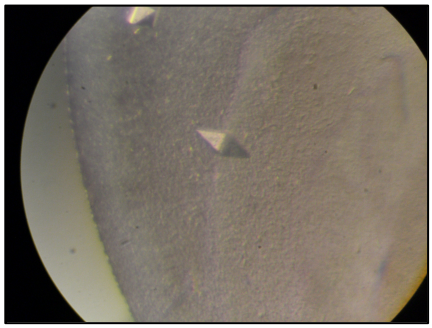
3.6 Subunit binding may cause conformational changes of FlhB

$FlhB_C$ was previously shown to form crystals that could be solved with atomic resolution using standard X-ray crystallography techniques (192). Three different buffer screens were created (Figure 2.2 and 2.3) to grow $FlhB_C$ with and without GRM peptide to determine the atomic contacts between early subunit GRM and the surface-exposed hydrophobic binding pocket. Screens GT2 and GT3 were optimised based on the success of protein crystallisation in screen GT1. Peptide was added at five-fold molar excess to protein to encourage co-crystallisation. Crystals were first observed after 1 week incubation.

Many wells showed extensive precipitation, but protein crystals similar to those previously reported were seen in the presence of all peptides used (Figure 3.11A) (237). Upon moving them to cryostorage, it was observed that the presence of the GRM peptide had weakened the $FlhB_C$ crystals compared to crystals in the absence of GRM peptide. Crystals were analysed by Dr. Philip Hinchliffe using beamline I24 at Diamond Light Source UK. Dr. Hinchliffe also performed the data analysis generated from X-ray diffraction and generated the PyMol structure of the resulting unbound $FlhB_C$ (Figure 3.11B).

Only one crystal returned a near-atomic resolution structure (at 4.0 Å). This crystal, grown in the presence of FlgE GRM peptide, did not show any additional density compared to the published structure for $FlhB_C$, suggesting that the FlgE GRM peptide had not co-crystallised with the protein. Since crystallography was not possible with

A



B

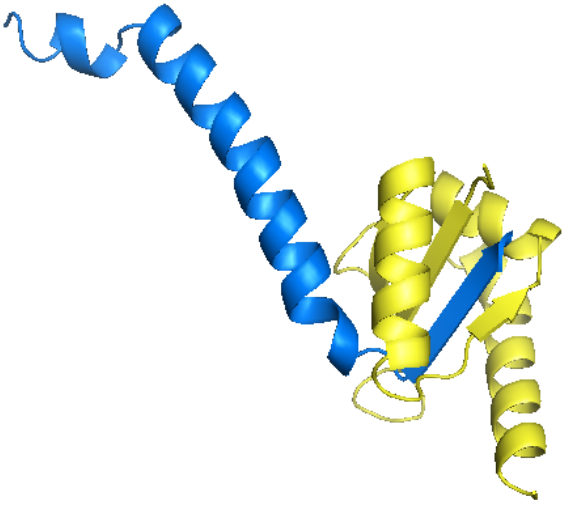


Figure 3.11 Crystallisation of $FlhB_C$

(A) $FlhB_C$ protein crystal in buffer containing a five-fold molar excess of FlgE GRM peptide at 10x magnification using a stereomicroscope.

(B) Near-atomic resolution structure of unbound $FlhB_C$ resolved to 4.0 Å. No additional election density was uncovered, suggesting the FlgE GRM peptide had not bound $FlhB_C$ during co-crystallisation. Dr. Philip Hinchliffe analysed the crystals using beamline I24 at Diamond Light Source UK and performed the molecular replacement to generate the structure.

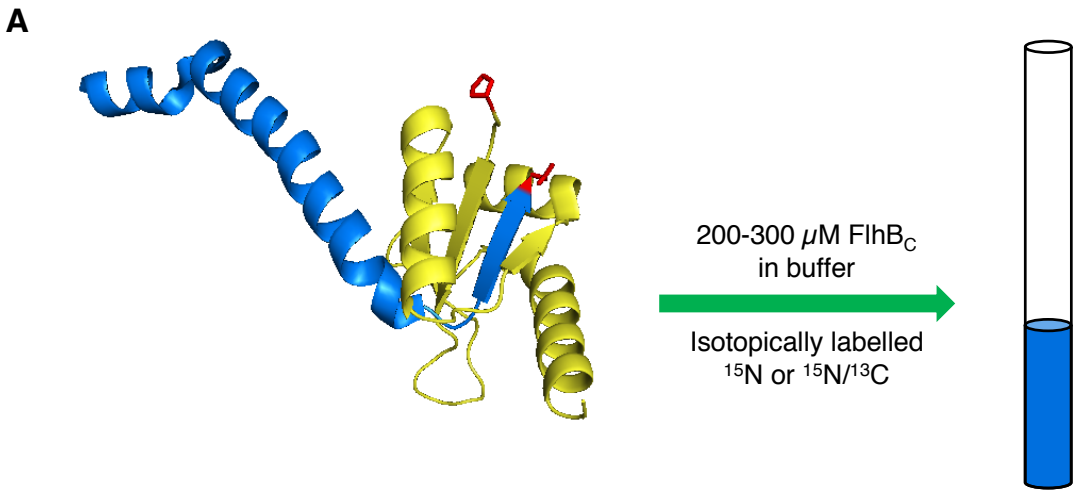
$FlhB_C$ and the GRM peptides, the decision was made to observe the effects of GRM peptide binding on $FlhB_C$ in solution as had been done with ITC.

NMR spectroscopy was used to characterise the effect of early subunit binding on $FlhB_C$ (Figure 3.12). This allowed both protein and peptide to be kept in solution in physiologically relevant buffer. NMR spectroscopy was conducted and analysed with the assistance of Dr. Daniel Nietlispach, who helped plan the experiments and collect and process the data, and Dr. Yvonne Yu, who assisted with assigning the peaks to the $FlhB_C$ sequence.

Briefly, isotopically labeled ($^{15}NH_4Cl$, Sigma) $FlhB_C$ was analysed using a Bruker 800MHz NMR spectrometer with a cryoprobe to detect proton and ^{15}N signals in the presence and absence of unlabeled peptide. Three peptides were chosen based on their final assembly location of their “parent” subunit in the flagellum and their affinity for the export gate. Additionally, the FlgE F39K GRM peptide variant (Figure 3.11B) and $FlhB_C$ containing a point mutation in the subunit binding pocket ($FlhB_C$ $P_{287}E$) (Figure 3.16) were used as non-binding controls.

To assign the peaks in the various spectra, nitrogen-15 and carbon-13 labeled $FlhB_C$ was studied using a Bruker 600MHz NMR spectrometer with a cryoprobe to detect proton, nitrogen-15, and carbon-13 signals. Peak assignment reached about 70% of total protein sequence, not including prolines as these residues are not visible in the nitrogen-15 spectrum because of the lack of a primary (1°) amide proton.

It was noted that over the course of the week-long 3D experiments, a secondary cleavage site in $FlhB_C$ between residues D_{237} and P_{238} resulted in the movement and loss of some peaks. A non-autocleavable mutant of $FlhB_C$, $P_{238}A$, was developed and labeled with deuterium, nitrogen-15, and carbon-13 (Figure 3.12B). This mutant still autocleaved at the primary, and biologically relevant, site between residues N_{269} and P_{270} (193). This deuterated, dual labeled (^{15}N and ^{13}C) sample allowed further peaks to be assigned to the protein sequence, about 80% of the protein not including the prolines (Figure 3.13), allowing for a comparison between peptide-bound and unbound $FlhB_C$.



2D sample (proton + ^{15}N) to measure subunit-bound and unbound FlhB$_C$ states.

3D sample (proton/^{2}H + ^{15}N + ^{13}C) to assign residues to signal peaks in spectra.

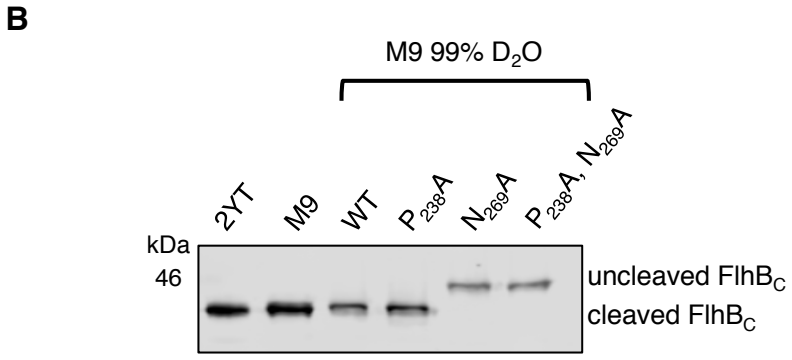
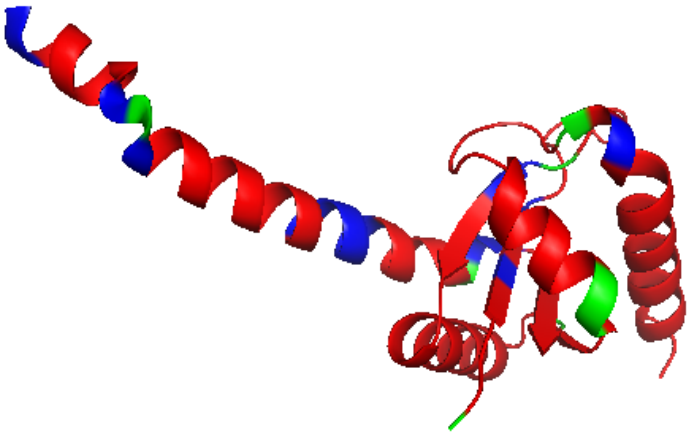


Figure 3.12 Overview of NMR experiments using FlhB$_C$ (residues 219-383, G$_{383}$A)

(A) Concentrated FlhB$_C$ in 50 mM Tris-HCl, pH 7.4 buffer with 150 mM NaCl was carefully pipetted into a thin-walled glass tube for use in a Bruker NMR spectrometer. Experiments were run at 25°C. FlhB$_{CN}$ is coloured blue, FlhB$_{CC}$ is coloured yellow, and the primary autocleavage site between residues N$_{269}$ and P$_{270}$ is coloured red with side groups shown. FlhB$_C$ was purified as described in Figure 3.3.

(B) Immunoblot of different conditions for FlhB$_C$ expression. Whole cell samples of expressed protein were analysed using mouse anti-CBD (NEB) and donkey anti-mouse infrared secondary antibody (Li-Cor). Samples were imaged using a Li-Cor Odyssey CLx and Image Studio Lite (Li-Cor). 2YT is wild type FlhB$_C$ expressed in 2x Yeast-Tryptone medium; M9 is wild type FlhB$_C$ expressed in M9 minimal medium. The other samples were expressed in M9 minimal medium with 99% D$_2$O. WT is wild type FlhB$_C$. N$_{269}$A and P$_{238}$A, N$_{269}$A are non-autocleavable mutants of FlhB$_C$.

A



B

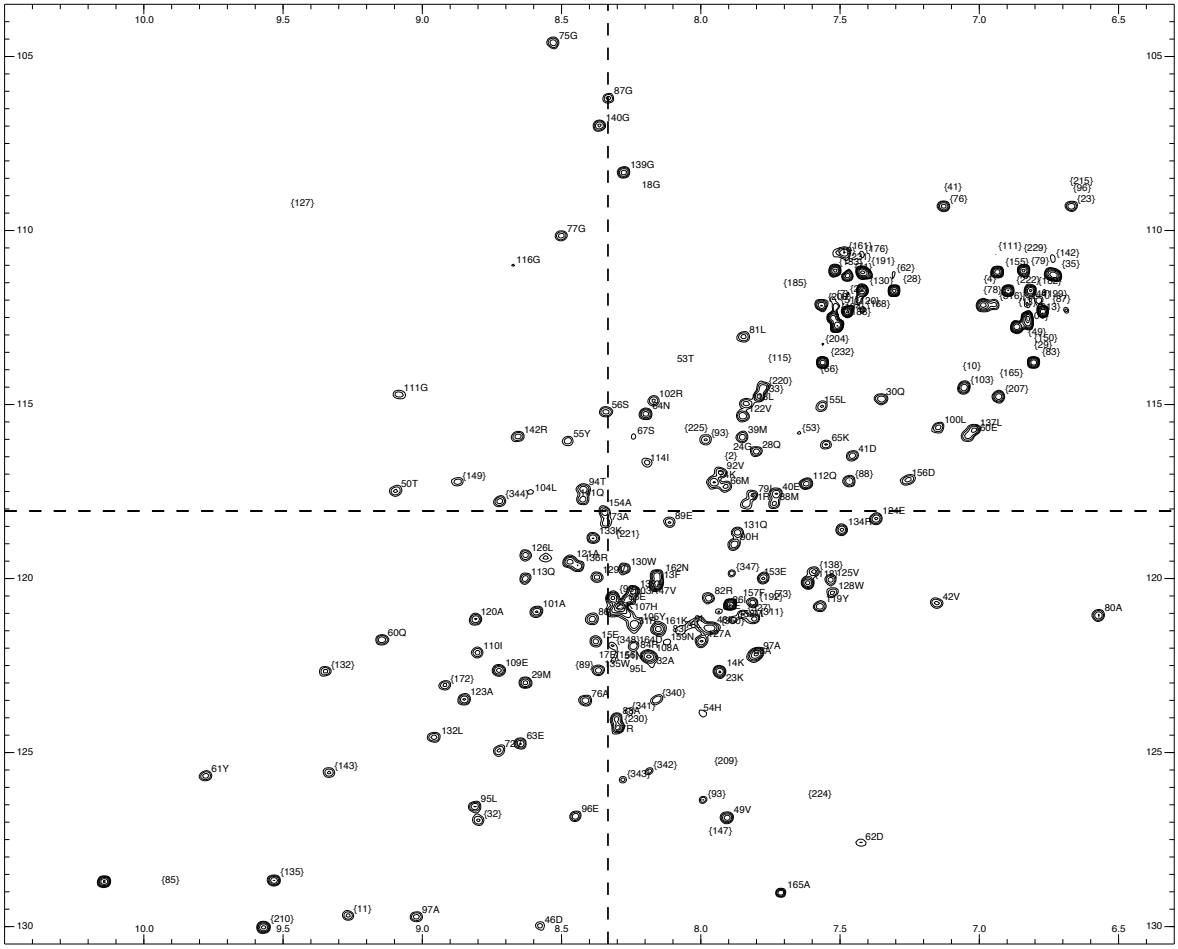


Figure 3.13 Assignment of $FlhB_C$ residues to peaks generated by NMR

(A) Peaks recorded by NMR spectroscopy were assigned to the primary sequence of $FlhB_C$ before being mapped onto the atomic resolution structure of $FlhB_C$. Red residues are assigned to peaks in the NMR spectra, blue residues could not be assigned, and green are prolines, which are not visible in proton and ^{15}N (2D) SOFAST-HMQC spectra. Residues R_{354}-R_{360}, E_{371}-N_{380}, and D_{382}-A_{383} are assigned but not represented above as they are not a part of the atomic resolution structure.

(B) 2D HMQC spectrum of unbound $FlhB_C$ with assigned residues labeled. X-axis is ^{1}H signal (ppm). Y-axis is ^{15}N signal (ppm).

Wild type $FlhB_C$ was titrated stepwise with GRM peptide at an initial concentration of 4.8 mM. Peptide was added in 50 μl increments until the $FlhB_C$ peaks no longer shifted. This corresponded to a final GRM peptide concentration of 550 μM. Titration with FlgE GRM peptide caused multiple chemical shift perturbations in $FlhB_C$ (Figure 3.14A), indicating either multiple binding events or a high degree of secondary effects caused by subunit peptide binding to $FlhB_C$. The FlgE F39K GRM peptide caused little change in $FlhB_C$, confirming the earlier ITC observation that the peptide cannot bind $FlhB_C$ *in vitro* (Figure 3.14B).

Similar to the FlgE GRM peptide, both the FlgB GRM peptide (Figure 3.15A) and the FliK GRM peptide (Figure 3.15B) caused a high number of chemical shift perturbations in $FlhB_C$. Although all three peptides caused shifts in similar residues, there were noticeable differences in the magnitudes and directions of the shifts. $FlhB_C$ $P_{287}E$ titrated with the FlgB GRM peptide showed few chemical shift perturbations across the protein (Figure 3.16).

Chemical shift perturbations in $FlhB_C$ were compared by assigning residues across spectra from peptide-bound and unbound protein. The chemical shift perturbations observed in the unbound state (the reference spectrum) were subtracted from the chemical shift perturbations in the GRM peptide-bound state. The difference in shifts (Δδ) was analysed to identify residues that showed the greatest shift upon GRM peptide binding (Figure 3.17).

For all wild type GRM peptides, FlhB $Alanine_{341}$ (A_{341}), a residue in the proposed early subunit binding pocket, showed the greatest chemical shift perturbations (over 0.4 ppm) (86). For comparison, most $FlhB_C$ residues experienced less than a 0.05 ppm chemical shift perturbation upon GRM peptide titration and binding. $Leucine_{344}$ (L_{344}), also in the proposed early subunit binding pocket, showed large chemical shift perturbations, although often not as great as other residues (e.g. $Isoleucine_{328}$ and $Alanine_{338}$, residues not in the proposed binding pocket). Another member of the proposed early subunit binding site, $Alanine_{286}$, showed only a small shift perturbation upon subunit binding. The final residue of the proposed early subunit binding site,

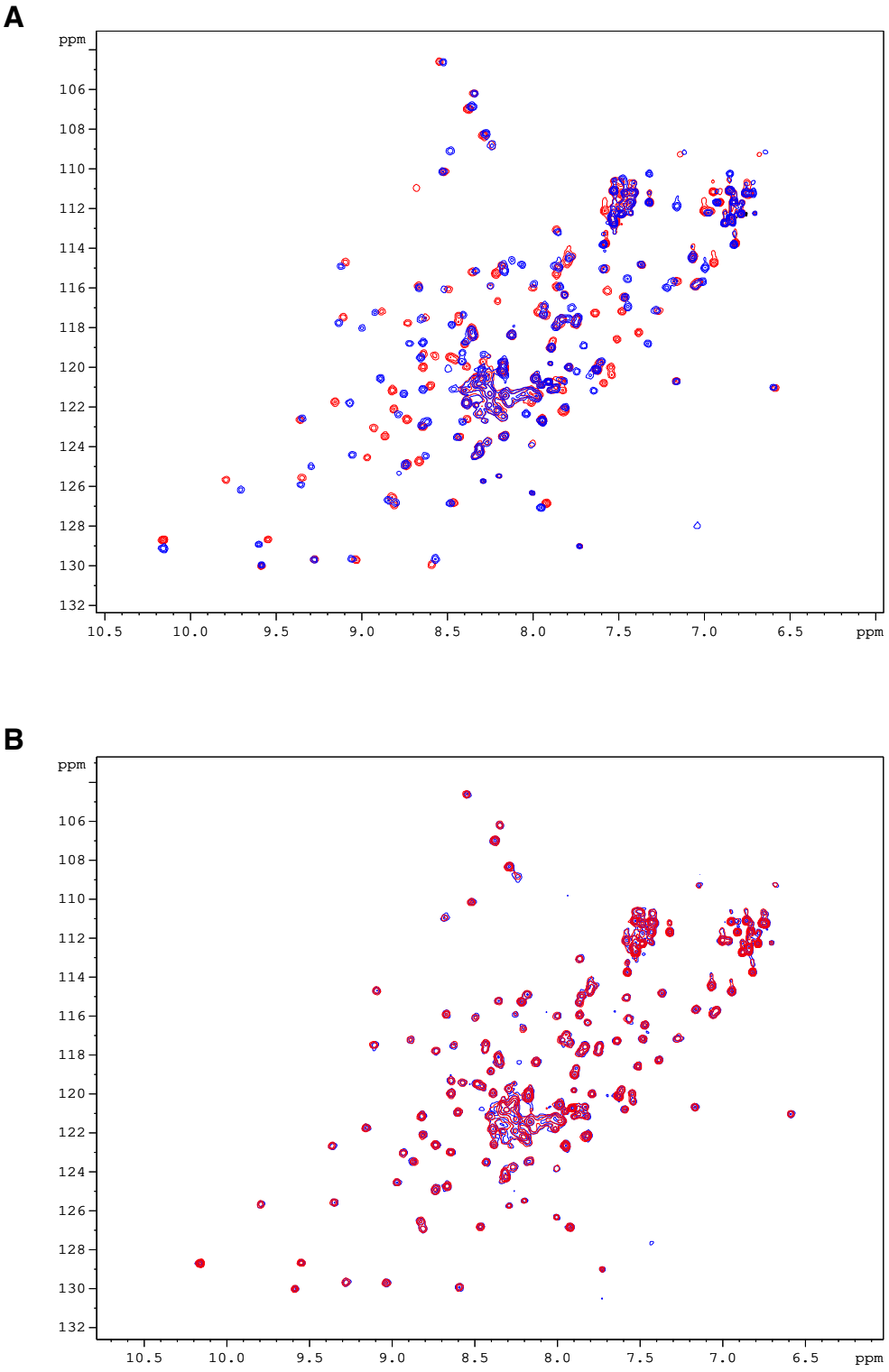


Figure 3.14 Wild type GRM peptides change the $FlhB_C$ environment

(A) Spectra overlay of $FlhB_C$ titration experiment with FlgE GRM peptide. Peaks in blue are fully titrated, or bound, $FlhB_C$, while peaks in red are unbound $FlhB_C$. The x-axis is the ^{1}H signal, the y-axis is the ^{15}N signal.

(B) Spectra overlay of $FlhB_C$ titration experiment with FlgE F39K GRM peptide. Peaks in blue are bound $FlhB_C$, while peaks in red are unbound $FlhB_C$. The x-axis is the ^{1}H signal, the y-axis is the ^{15}N signal.

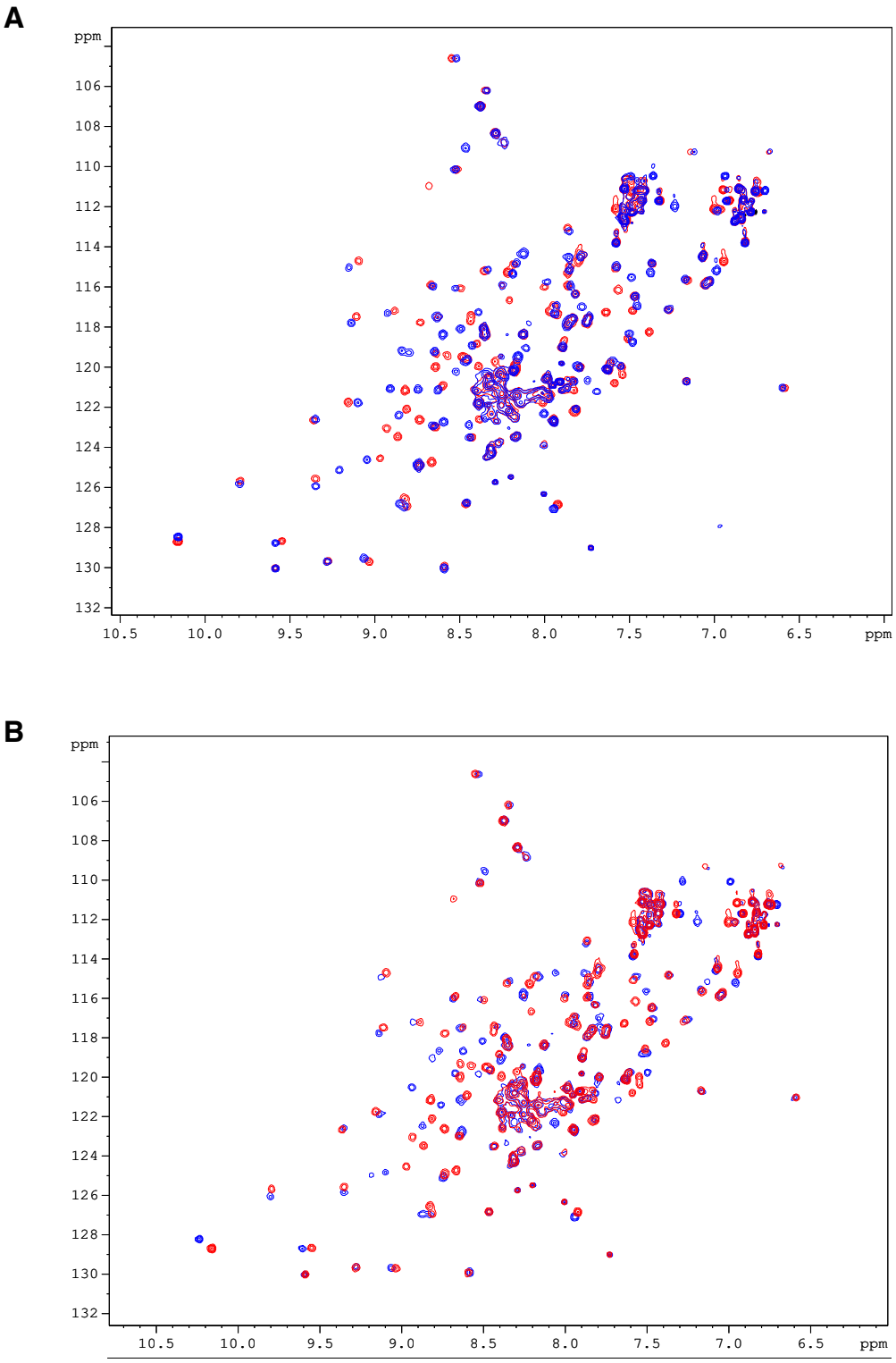


Figure 3.15 Wild type GRM peptides change the $FlhB_C$ environment

(A) Spectra overlay of $FlhB_C$ titration experiment with FlgB GRM peptide. Peaks in blue are fully titrated, or bound, $FlhB_C$, while peaks in red are unbound $FlhB_C$. The x-axis is the ^{1}H signal, the y-axis is the ^{15}N signal.

(B) Spectra overlay of $FlhB_C$ titration experiment with FliK GRM peptide. Peaks in blue are bound $FlhB_C$, while peaks in red are unbound $FlhB_C$. The x-axis is the ^{1}H signal, the y-axis is the ^{15}N signal.

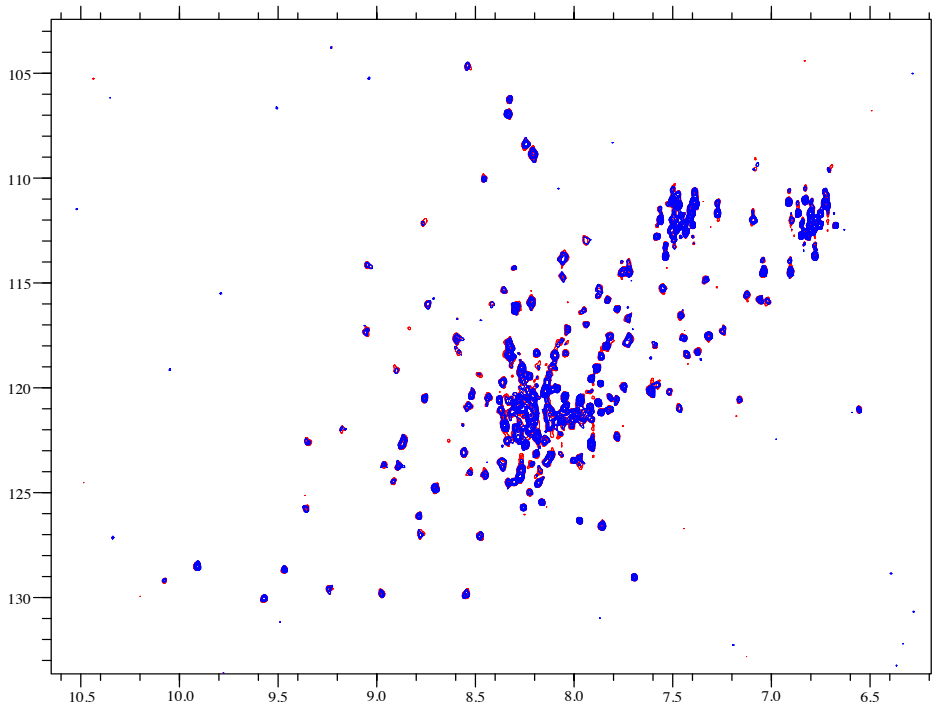


Figure 3.16 Mutating the proposed GRM binding pocket on $FlhB_C$ disrupts subunit-$FlhB_C$ interactions

Spectra overlay of $FlhB_C$ $P_{287}E$ titration experiment with FlgB GRM peptide. Peaks in red are fully titrated, or bound, $FlhB_C$ $P_{287}E$, while peaks in blue are unbound $FlhB_C$ $P_{287}E$. The x-axis is the ^{1}H signal in ppm, the y-axis is the ^{15}N signal in ppm.

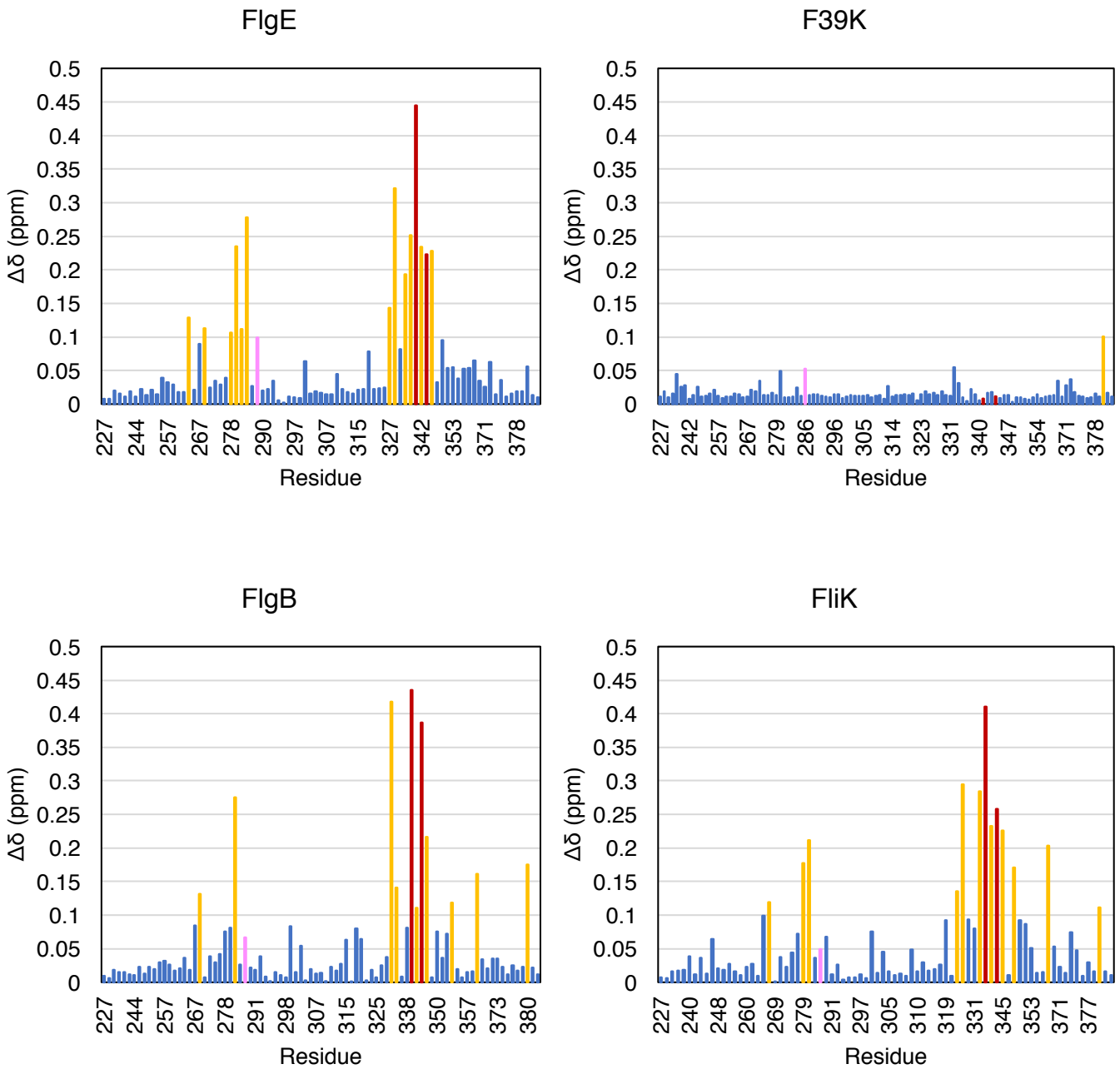


Figure 3.17 Chemical shift perturbations in the ^{1}H and ^{15}N resonances experienced by $FlhB_C$ residues when bound by GRM peptide

The title of each plot represents the GRM peptide used in that titration experiment. Peaks of the proposed early subunit binding site are in red and purple, peaks experiencing greater than 0.1 ppm shift when bound by peptide are in gold. The chemical shifts (Δδ) are compound chemical shifts for each backbone ^{1}H-^{15}N signal calculated as described in materials and methods.

Proline$_{287}$, could not be detected by 2D NMR, so a chemical shift perturbation pattern was not determined.

Residues near the proposed binding pocket also showed chemical shift perturbations greater than 0.1 ppm, but the perturbations are not limited to residues close to the subunit binding pocket in FlhB$_C$. Two residues, Glutamate$_{327}$ (E$_{327}$) and Isoleucine$_{328}$ (I$_{328}$), are located on the loop between FlhB$_C$ α-helices 3 and 4, on the face opposite the proposed binding pocket showed perturbations greater than 0.1 ppm. I$_{328}$ showed chemical shift perturbations regardless of the peptide present, whereas E$_{327}$ only showed shift perturbations upon binding with FliK and FlgE GRM peptide. Tyrosine$_{268}$, a residue located on FlhB$_{CN}$ just before the autocleavage site, also showed chemical shift perturbations upon peptide binding, regardless of subunit used. Residues experiencing chemical shift perturbations greater than 0.1ppm were mapped onto the atomic resolution structure (Figure 3.18).

FlhB$_C$ peak assignments in the peptide-bound state account for only 85-90% of known residue peaks assigned in the unbound reference FlhB$_C$ spectrum. Therefore, additional chemical shift perturbations near the proposed binding site or elsewhere on the protein could be possible. The majority of known chemical shift perturbations occurred around the proposed subunit binding site, particularly residues A$_{341}$ and L$_{344}$, and any further perturbations could reasonably be expected to be localised to the proposed binding site and surrounding area. So, it is likely that the data presented thus far provides robust coverage of subunit GRM peptide binding to FlhB$_C$ and confirms the FlhB$_C$ surface-exposed hydrophobic patch as the binding site for early subunits.

A sequence alignment of FlhB and its homologues was conducted to determine how conserved the hydrophobic binding pocket might be (Figure 3.19). Alanine$_{286}$, which did not experience great chemical shift perturbations upon early subunit binding, is the only binding pocket residue not well conserved, being replaced with leucine and valine in some species. Proline$_{287}$ is found in all species and homologues sampled, including the FliR-FlhB fusion protein used by *Clostridium difficile*, but is invisible in 2D NMR and difficult to examine using this method. A$_{341}$ is found in all species and homologues

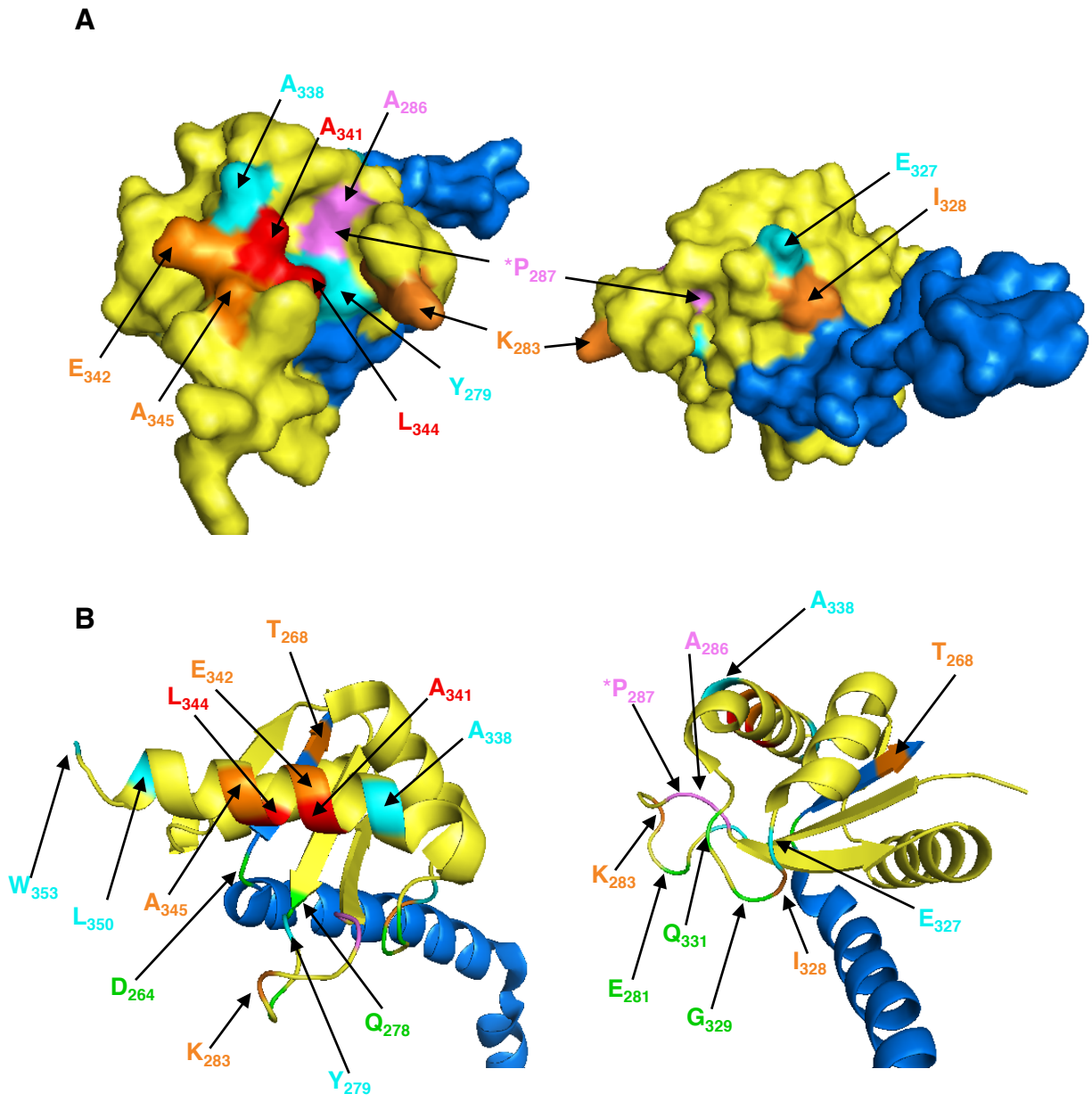


Figure 3.18 Residues that experienced chemical shift perturbations greater than 0.1 ppm in the ^{1}H and ^{15}N resonances upon subunit binding are mapped and labeled on the $FlhB_C$ atomic resolution structure

(A) Surface view. $FlhB_{CN}$ is coloured blue, $FlhB_{CC}$ is coloured yellow. The surface-exposed hydrophobic subunit binding pocket residues that experienced chemical shift perturbations greater than 0.1 ppm are coloured red, while the other binding pocket residues are coloured purple. Orange residues experienced chemical shift perturbations greater than 0.1 ppm in all three titration experiments; cyan residues experienced chemical shift perturbations greater than 0.1 ppm in two of the three titration experiments. $*P_{287}$ does not have an amide proton and could not be detected in 2D NMR titration experiments.

(B) Ribbon view. $FlhB_{CN}$ is coloured blue, $FlhB_{CC}$ is coloured yellow. The surface-exposed hydrophobic subunit binding pocket residues that experienced chemical shift perturbations greater than 0.1 ppm are coloured red, while the other binding pocket residues are coloured purple. Orange residues experienced chemical shift perturbations greater than 0.1 ppm in all three titration experiments; cyan residues experienced chemical shift perturbations greater than 0.1 ppm in two of the three titration experiments; and green residues experienced chemical shift perturbations greater than 0.1ppm in one of the three titration experiments.

Figure 3.19 Sequence alignment of export gate component, FlhB, from selected species

The early subunit surface-exposed hydrophobic binding pocket of *S. enterica* is boxed, showing conservation in other species and across T3SS. Green residues experienced chemical shift perturbations when titrated with one peptide; blue residues experienced chemical shift perturbations when titrated with two peptide; and orange residues experienced chemical shift perturbations when titrated with three peptides. Only N_{380} experienced significant chemical shift perturbations when titrated with the FlgE $F_{39}K$ peptide.

The FliR-FlhB fusion protein from *Clostridium difficile* has been truncated to remove most of the FliR domain, which did not align to the other sequences. YscU and PscU are FlhB homologues found in the injectisome. Sequences were acquired from UniProt and aligned with Clustal Omega using default settings. *Treponema pallidum* (strain Nichols); *Borrelia burgdorferi* strain ATCC 35210; *Yersinia pestis*; *Pseudomonas aeruginosa* (strain ATCC 15692); *Helicobacter pylori* (strain ATCC 700392); *Escherichia coli* strain K12; *Salmonella enterica* serovar Typhimurium strain LT2; *Aquifex aeolicus* strain VF5; *Bacillus subtilis* strain 168; *Clostridium difficile*.

save for the FlhB found in the spirochetes. L_{344} is always a larger hydrophobic residue, sometimes being replaced with phenylalanine. This sequence alignment, along with recent findings in the injectisome T3SS, strongly suggests that this surface-exposed, hydrophobic early subunit binding pocket is conserved across all T3SSs (88, 89).

3.7 Discussion

The data described in this chapter highlight the importance of the early subunits' GRM, especially the highly-conserved phenylalanine, and the environmental changes caused by early subunit binding to $FlhB_C$.

In vivo data investigating the deletion or mutation of the FlgE GRM (e.g. $F_{39}K$, $F_{43}K$ labeled as FFKK in Figures 3.1 and 3.2) on subunit export and cellular motility suggest that the first and fifth residues of the GRM are the most important residues for early subunit binding to $FlhB_C$ and eventual subunit export. This isn't surprising as these are the most conserved residues in the early subunit GRM across subunits and species (this study and 85). The middle three GRM residues, although not investigated in this study, vary greatly across subunits and species. What effects they may have on subunit export and cellular motility are currently unknown. I hypothesise that these residues are most important for the correct positioning of the first and fifth GRM residues when early subunit binds to $FlhB_C$. Further experimentation on these residues may be necessary to determine structure of the GRM when bound to $FlhB_C$ and how these middle residues affect early subunit binding.

ITC and NMR data corroborate the *in vivo* data on the importance of the FlgE GRM as the single replacement mutants (e.g. $F_{39}K$ and $F_{43}K$) on synthetic GRM peptides were unable to bind to $FlhB_C$ or cause the chemical peak shifts seen with wild type peptide. None of the *flgE* GRM deletions or mutations were dominant when expressed in wild type *Salmonella* and assayed for swimming motility, suggesting that wild type FlgE can outcompete the FlgE GRM deletions or mutations *in vivo* probably due to their lack of strong binding to $FlhB_C$.

All early subunits binding affinities to $FlhB_C$ fall in the micromolar range, confirming that the interaction between $FlhB_C$ and the early subunits is transient. These affinities corroborate the previously observed binding affinity of early subunits and $FlhB_C$ (86,

191). The FlgD and FliK GRM peptides bound $FlhB_C$ with similar affinities to full length FlgD or FliK, indicating that the subunit GRM peptides encompass the major binding site for $FlhB_C$ and that the GRM is primarily responsible for early subunit binding affinity to $FlhB_C$.

Not all the ITC experiments reached full saturation. Additionally, a few of the curves show unusual spikes or dips in the data. Some of the GRM peptides are difficult to maintain in solution due to the hydrophobicity of the GRM and surrounding sequence, so it is possible that some of the peptide aggregated and precipitated during ITC experiments. This would lower the concentration of peptide available for titration and impact the quality curves. All ITC experiments were repeated at least three times under the same conditions. Additional experiments may be necessary to take the peptide-$FlhB_C$ interaction to saturation and confirm the magnitude of the micromolar affinities reported here.

The transient interaction between the export gate and early subunit may be important to keep subunits moving to the distal site of assembly, whether by a subunit chain or diffusion. If subunits were to dock tightly to FlhB prior to export, it would require more energy to pull subunits into the chain or require more energy to inject them into the channel (86, 99). This could slow subunit export and, ultimately, flagellum growth. Observations showing that flagellar growth rate is constant, regardless of length, have been challenged by other work showing that growth rate is length-dependent (99, 172). These observations were conducted on the long filament and it is unknown whether the rate of HBB growth is constant or variable. The HBB structure is only about 100 nm in length, or the length of a fully stretched FlgE subunit, so it is unlikely that the subunit chain or diffusion growth rate is severely impacted by HBB length. Still, subunits strongly binding to the export gate would raise the energy barrier needed to export them, a costly enterprise for the cell.

Why might early subunit GRM binding affinities for FlhB differ? There is a general trend between subunit export order and subunit GRM binding affinity to $FlhB_C$, but, as the range is less than an order of magnitude, the correlation between subunit export order and binding affinity to $FlhB_C$ is weak. The binding affinities for the early subunit

GRM to $FlhB_C$ may also be a result of the final stoichiometry of each subunit in the early HBB structures. FlgE, the monomer, has one of the stronger binding affinities, and could outcompete FlgD and FliK to selectively bind FlhB to promote its own export. FliE also has a strong binding affinity for FlhB and must assemble before other rod subunits can. FlgB, the proximal rod subunit that assembles on FliE, shares a similarly strong binding affinity for FlhB.

More work is needed to fully understand the role of early subunit GRM binding affinities for FlhB. Kinetics of early subunit-$FlhB_C$ binding, such as the on and off rates for early subunit binding, can be studied to determine how subunit-gate binding may influence subunit export order, if at all. NMR could be used to measure the dynamics of early subunit binding to $FlhB_C$ by observing the rate at which peaks shift and broaden in the GRM peptide-bound versus unbound $FlhB_C$ state. Other biophysical techniques could also be used to establish the reaction parameters, k_{on} and k_{off}, which may play a greater role than binding affinity in subunit export order. Finally, expression levels of each subunit prior to export are needed to put these affinities and reaction kinetics into context. Weaker binders may be better expressed and could then outcompete the stronger binders, for example, but it is difficult to conclude this based on available proteomic data.

Intriguingly, FlgC is the only *Salmonella* early subunit to contain an alanine in the fifth GRM position, rather than a large hydrophobic residue. Although an affinity for FlgC GRM peptide to $FlhB_C$ could not be established using ITC, mutational analysis and swimming motility assays confirm the importance of the conserved phenylalanine in FlgC's GRM. Mutating the conserved phenylalanine in the FlgC GRM creates a GRM similar to the previously described FlgE mutant FFAA, which was shown to have impaired export and to swim poorly when compared to wild type FlgE. Swapping the *flgC* GRM for the native GRM sequence from *flgB* did restore motility in a *flgC* null strain. As this GRM swap was not dominant in wild type *Salmonella*, this mutant adds further evidence that it is the presence of the first phenylalanine in the GRM that is of vital importance.

The addition of DMSO to the buffer did help with FlgC GRM peptide solubility, but still no binding to $FlhB_C$ could be observed. This suggests the binding affinity for the FlgC GRM to $FlhB_C$ may be too weak to be measured by ITC. It is also possible that DMSO interfered with the putative hydrophobic interaction between the FlgC GRM peptide and $FlhB_C$, but an ITC experiment using $FlhB_C$ and FlgB GRM peptide showed only a twofold increase in the binding affinity in the presence of DMSO (Figure 3.8). Local concentration of FlgC near the export machinery may be high enough to force an interaction between FlgC and $FlhB_C$ or other rod subunits may act as partner proteins for FlgC export, requiring less interaction between the FlgC GRM and $FlhB_C$. *In vitro* work in this study between FlgB and FlgC suggests that these two rod subunits do not interact. This work did use an FPLC and a low flow rate to test interactions. Column pressure was closely monitored, but it is possible any interaction between FlgC and other rod subunits is too weak to be observed by co-purification using FPLC. Follow up work, such as protein pulldown assays that examine potential interactions, is needed.

A sequence alignment of FlgC from seven different bacterial species emphasises the unusual nature of this GRM when compared to other subunits such as FlgE (Figure 3.9A). The *Salmonella* FlgC GRM starts at residue 49, further downstream than in other subunits. The GRM is defined as an N-terminal secretion motif, but, in a protein of 134 amino acids, FlgC's GRM is more centrally located than the GRM in FlgE or FliK. FlgB and FlgC are reported to be completely disordered *in* vitro, suggesting they may only adopt a fold once assembled into the flagellar rod (132). So, although the conserved phenylalanine occurs more centrally in FlgC, it will still be disordered and bind to $FlhB_C$ in preparation for export.

The Gammaproteobacteria in the sequence alignment all have an alanine in the fifth GRM position, while the FlgC GRM in other species is either a leucine or valine Experiments using FlgC from *Bacillus*, *Agrobacterium*, or *Borrelia* species could be conducted to see if their variants with a more typical residue in the fifth position bind to $FlhB_C$. If that is the case, it is possible the Gammaproteobacteria evolved other mechanisms for the export of FlgC, as discussed previously. *Aquifex* and *Borellia* are unusual in that they contain another phenylalanine downstream of the conserved FlgC

GRM phenylalanine. The role this downstream phenylalanine may play in subunit export is unknown.

Prior research is mixed on whether the rod proteins are dominant when overexpressed in wild type *Salmonella*. Minamino and colleagues saw no effect on wild type motility when overexpressing *flgB*, but did report a slight effect when expressing a mutant of *fliE* at 30°C (131). Hirano and colleagues reported an overexpression effect from all the rod proteins, including *flgB*, when expressed in wild type *Salmonella* (238), although *flgB* and *flgC* had less of an effect compared to other rod subunits. In both cases, the studies used pTrc99A, with the latter study using a further optimised plasmid called pTrc99AET (238).

pTrc99AET is isopropyl β-D-1-thiogalactopyranoside (IPTG) inducible, whereas the pBAD18amp plasmid used in this study to express *flgC* in *Salmonella* is inducible by L-arabinose. Siegele and Hu report that the addition of 1.33 mM L-arabinose, or 0.2% w/v arabinose, was enough to cause 89% of cells to express GFP, as rapidly as 15 minutes post-induction, and concluding that the dark cells had lost the GFP plasmid (239). Starter cultures, used to inoculate the swimming motility agar, contained wild type *Salmonella* carrying the pBAD18amp plasmid with and without the *flgC* GRM variants were first incubated to an A_{600} of 1.0 in the presence of 0.02% w/v arabinose to induce protein expression. Therefore, the pBAD18amp plasmids should have already been producing protein before inoculation on agar containing the higher arabinose concentration. It is then likely that the differences in overexpression can be explained by the difference of induction systems used.

Further work is clearly needed to unpick the importance of the phenylalanine residue to export or assembly of FlgC and the impact of the alanine in the fifth GRM position. *In vivo* experiments, such as cell fractionations after osmotic shock, could be used to study the ultimate fate of the FlgC GRM mutant, F49A, as export or assembly deficient to establish the importance of the phenylalanine. Furthermore, in the native FlgC GRM, the fifth position alanine may be insufficient to establish contacts with the binding pocket on $FlhB_C$, allowing only for the phenylalanine to contact the binding pocket, if the phenylalanine were required for export. The middle three residues of the

FlgC GRM may also play a greater role in navigating the alanine to the binding pocket or keeping the phenylalanine in closer proximity to the binding pocket.

Co-crystallisation of GRM peptide with $FlhB_C$ was unsuccessful, only returning a near-atomic resolution structures of unbound $FlhB_C$. We did observe a minor difference in the atomic resolution structure of $FlhB_C$ from the published structure. A part of the $FlhB_{CN}$ α-helix presented as a random coil in the structure presented in this chapter (Figure 3.11B). Because our structure could only be resolved to 4.0 Å, while the published structure was resolved to 2.45 Å, this minor structural difference is most likely due to the poorer resolution in our analysis than caused by any peptide binding (192).

Co-crystallisation was attempted concurrently with the ITC experiments, so affinity measurements between subunit GRM peptides and $FlhB_C$ were not completed when the crystal screens were conducted. Later screens were attempted that decreased the molar excess of peptide to $FlhB_C$ while increasing the overall $FlhB_C$ concentration but these failed to crystallise. The transient interactions observed via ITC may be too weak to accommodate co-crystallisation between peptide and $FlhB_C$. Further experiments could see unbound $FlhB_C$ protein crystals soaked in peptide and cryoprotectant prior to freezing. Crystallisation with the early subunit peptides covalently fused to $FlhB_C$ or co-crystallisation between $FlhB_C$ and peptide fused to a partner protein are also possibilities given the ability to purify $FlhB_C$ in high yields.

The use of NMR to study subunit binding proved fruitful, although there were complications along the way. The $FlhB_C$ 3D spectra recorded have fairly discrete peaks and it looked easy to assign residues to the spectra. Assignments took longer than anticipated, however, for several reasons. Firstly, a region of overlapping peaks proved to be a central part of $FlhB_C$, rather than the disordered termini. As these peaks overlapped, it is possible some residues were masked by other, stronger signals. This may also explain why certain discrete peaks could not be assigned to the $FlhB_C$ sequence as masked residues could disrupt sequential assignment. Secondly, the high number of prolines in the protein caused some difficulty. Because prolines do not possess an amide bond in a polypeptide, they do not produce an amide proton signal

and are therefore hidden in 2D NMR experiments using a probe for ^{1}H and ^{15}N. $FlhB_C$ has 13 prolines, roughly 8% of the protein. This makes sequential assignment of peaks to the amino acid sequence difficult in proline-rich regions. Deuterated $FlhB_C$ was used to push assignments further than proton-labeled $FlhB_C$ allowed. The deuterated sample did allow peak assignments covering roughly 80% of the protein, not including prolines, enabling robust comparison across the bound and unbound $FlhB_C$ states.

The proposed binding pocket identified by Evans and colleagues was corroborated by NMR in this study (86). The FlhB sequence alignment confirms that this binding pocket is well conserved across species. $Proline_{287}$ of the binding pocket is well conserved across all species and homologues sampled. Although this residue could not be seen in the 2D NMR experiments, and so its contribution to subunit binding is left undetermined, its conservation suggests that it is important to at least the makeup of the binding pocket.

$Alanine_{286}$, one of the four proposed residues and the least well conserved, did not experience a large chemical shift perturbation upon GRM peptide binding. It is unlikely, then, that this residue interacts with early subunits or is altered once subunit binds. This alanine could be merely a structural feature to form the surface-exposed hydrophobic binding pocket on $FlhB_C$'s surface.

$Alanine_{341}$ and $Leucine_{344}$, both well conserved residues, appear to be involved directly in early subunit binding. $Alanine_{341}$ experiences the greatest chemical shift perturbations in the presence of all three GRM peptides, with $Leucine_{344}$ experiencing smaller perturbations. Combining the data from the NMR titration data with from the early subunit export and motility assays suggests that $Alanine_{341}$ directly interacts with the conserved phenylalanine of early subunit GRM. $Leucine_{344}$ may also interact with the phenylalanine, interact with the fifth hydrophobic GRM residue, or serve to stabilise the interaction between $Alanine_{341}$ and the early subunit GRM phenylalanine.

Surrounding residues, such as $Lysine_{283}$, may move away from the binding pocket upon early subunit binding as the hydrophobic residues would no longer need to be shielded from the hydrophilic environment. $Lysine_{283}$ could also act as a clamp, of

sorts, to hold subunit into the binding pocket via a salt bridge or polar interaction with residues near the GRM. Residues more distant from the binding pocket probably experience chemical shift perturbations as a secondary effect of subunit binding causing rearrangement to the binding pocket structure. As both $Glutamate_{327}$ and $Isoleucine_{328}$ are located on the loop immediately prior to α-$helix_4$, their chemical shift perturbations may result as the loop is repositioned to allow the binding pocket to accommodate early subunit.

That a tyrosine located on $FlhB_{CN}$ also experienced chemical shift perturbations is interesting. No other residues on $FlhB_{CN}$ experienced such perturbations, nor did any near the autocleavage site on $FlhB_{CC}$. The tyrosine's side chain does not protrude toward the binding pocket in the atomic resolution structure. In fact, the side chain is surface-exposed on another face of $FlhB_C$ distant from the binding pocket. Does this tyrosine play a greater role in early subunit recognition and binding? Or does this tyrosine need to move to accommodate early subunit binding to the binding pocket? Future work could investigate why a residue on $FlhB_{CN}$ was strongly affected by early subunit binding.

Residues not found in the published atomic resolution structure did experience chemical shift perturbations upon binding in the presence of FlgB and FliK GRM peptide. These residues are primarily located in the last 30 residues of $FlhB_C$, the C-terminal disordered tail region. Current NMR data are not robust enough to calculate the location or fold of these residues. Future work could seek to determine the structure and relevance of these disordered residues to subunit export.

Another line of inquiry could be to replicate these experiments using the non-autocleavable, early-locked gate, $FlhB_C$ $N_{269}A$. This variant of $FlhB_C$ can interact with early subunits, but does not allow for the switch in assembly phases. Does it share similar features to wild type FlhB? Or does autocleavage give FlhB the needed flexibility to properly recognise and interact with early subunits?

We can now conclude that early subunits interact transiently with $FlhB_C$ at a surface-exposed hydrophobic binding pocket, providing strong evidence that the hydrophobic

residues on $FlhB_{CC}$ directly interact with the GRM. But further work is needed to understand this fully. For instance, how does the conserved phenylalanine interact with the conserved binding pocket on $FlhB_C$, particularly $Alanine_{341}$ and $Leucine_{344}$? Do residues experiencing large chemical shift perturbations move out of the way to allow the protein to accommodate subunit? What do the differences in magnitude and direction of peak shifts mean? Do early subunits cause slightly different conformations in $FlhB_C$ and is this a result of accommodating each subunit or may this impose export order on the early subunits?

Work presented thus far has studied the interactions between early subunits and $FlhB_C$. But this is only one of FlhB's roles in controlling flagellar subunit export and assembly. One early subunit, FliK, is more than an exported early subunit. In tandem, FliK and FlhB are responsible for the export specificity switch at the fT3SS that leads to late subunit expression and export. The next chapter will explore this vital role of FlhB and seeks to provide clarity to the as of yet unsolved mechanism underlying the switch.

Chapter 4

The loss of a domain activates the export specificity switch at FlhB

4.1 Introduction

Much of the work investigating the export specificity switch concentrates on two key elements – the secretion of $FliK_N$ and the autocleavage of $FlhB_C$ (105, 126). Fewer studies have investigated why the autocleavage of $FlhB_C$ is important to the switching mechanism and how $FliK_C$ and $FlhB_C$ may interact to cause the switch.

Mizuno and colleagues found an acidic loop conserved across FliK and its SctP homologues. Their modeling suggests that the acidic residues of the loop are only able to access basic residues on the surface of $FlhB_C$ post-autocleavage and they proposed that these loops trigger the switch at $FlhB_C$ (203). Their hypothesis went untested, however, until follow-up work from the Aizawa group where they identified other residues in the loop, Leucine-Arginine-Leucine, that they suggest are important to the switch. Using a combination of electron microscopy and motility assays, they concluded that these three residues upstream of loop 2 were important to the switch by stabilising the core folded domain of $FliK_C$ where the T3S4 is located and not by interacting with $FlhB_C$ (240).

Currently, genetic evidence is strongest for the role of $FliK_C$ in triggering the switch. Scanning deletions of ten residues and truncations of $FliK_C$ resulted in a failure of the fT3SS to switch properly, leading to the polyhook or polyhook-filament phenotype (202, 204). Structural data was not available at the time, but now it is clear that these mutations remove large chunks of structured domains. Their deletion possibly destablises the rest of the C-terminal structure, again pointing to the importance of structural stability over any residue in particular.

In the *Yersinia* injectisome, chemical shift perturbations in the early subunit export gate were only observed by NMR spectroscopy upon titration with the molecular ruler disordered linker and C-terminal T3S4 domains. The $SctP_C$ T3S4 domain on its own did not cause any chemical shift perturbations in $SctU_C$, hinting that there is no direct

interaction between the T3S4 and export gate (89). Rather than suggesting a direct interaction between $FliK_C$ and $FlhB_C$ to trigger the switch, these data in combination strongly suggest a role of the specific structure of $FliK_C$ in triggering the switch, although it is unclear how.

McMurry and colleagues investigated the binding kinetics of FliK to wild type and non-autocleavable $FlhB_C$. They found that FliK bound to $FlhB_C$ with micromolar affinity (188). This agrees with my ITC data presented in the previous chapter which shows a micromolar affinity between FliK's GRM and $FlhB_C$. Their reported affinity was twofold stronger (K_d ~15 μM vs K_d ~8 μM) than the affinity I measured. The difference in affinities could be due to a difference in technique (ITC vs surface plasmon resonance) or their use of full length FliK instead of short peptides meant to study the GRM in isolation (188).

The lack of biochemical data on interactions between $FliK_C$ and $FlhB_C$ is not surprising. FliK/SctP has a relatively disordered N-terminus and linker region that could not be studied with techniques such as NMR and no atomic level resolution structure exists for these domains. The structured, "ball-like" $FliK_C$ domain only represents about 20% of the protein (203). This makes FliK a difficult protein to study. $FlhB_C$, on the other hand, is quite stable and easier to study.

The structure of autocleaved $FlhB_C$ revealed a core of four β-strands surrounded by four α-helices. The conserved NPTH autocleavage loop sits between two β-strands, β_1 and β_2 (192). These four β-strands seem to form a β-sheet that holds the two polypeptides together post-autocleavage with stabilising contributions from core-facing residues of the α-helices. Disrupting these interactions and destablising the structure of $FlhB_C$ is enough to restore functionality in an export specificity switch-deficient FlhB variant (241). Other suppressor mutations in $FlhB_C$ that partially overcome the effective loss of FliK were previously identified (242, 243). In all cases, these mutations were found on the C-terminal cytoplasmic domain of the early subunit export gate ($FlhB_{CC}$/$SctU_{CC}$). Most mutations resulted in a radical change of residue side chain biochemistry, such as a charged side chain becoming an aromatic hydrophobic side chain, although truncations of $FlhB_{CC}$ were also observed (242, 243).

These mutations were mapped onto the *Yersinia* early subunit export gate, SctU (YscU), and were able to partially restore effector molecule secretion (244). Accordingly, we have dubbed these mutations the “autonomous switchers” for their ability to trigger the export specificity switch in the absence of the molecular ruler.

Notably, these early subunit export gate mutants are more stable than the wild type protein, resulting in less efficient autocleavage (190). This reduced rate of autocleavage could explain why these mutants are only able to partially restore the wild type switching phenotype in the absence of the molecular ruler, although further work in this area is needed. These mutations are valuable tools to study the interactions between $FlhB_{CN}$ and $FlhB_{CC}$ and, ultimately, the export specificity switch in isolation from FliK.

Data presented in this chapter describes the role of FlhB autocleavage in the export specificity switch and suggest the switch is only triggered after a major structural change at $FlhB_C$. I will also discuss, briefly, how FliK may cause this change.

4.2 $FlhB_{CN}$ does not adopt a native fold in solution

To investigate how non-covalent interactions stabilise autocleaved $FlhB_C$, recombinant plasmids carrying $FlhB_{CN}$ and $FlhB_{CC}$ C-terminally fused to the intein-chitin binding domain (inteinCBD) were constructed. Protein was purified as described in the Materials and Methods.

Although a higher molecular weight band than expected for background inteinCBD was seen with immunoblotting, the fusion protein was too small to be full length $FlhB_{CC}$. Only inteinCBD could be detected by immunoblotting and staining with Coomassie R250 post DTT-mediated cleavage (Figure 4.1A). This expression was repeated after the plasmid was reconfirmed by sequencing, but $FlhB_{CC}$ was not recovered.

In contrast, $FlhB_{CN}$ expressed well and could be purified (Figure 4.1B). Before further attempts to purify $FlhB_{CC}$ were made, $FlhB_{CN}$ was analysed using circular dichroism (CD). The protein was scanned from 260 nm to 200 nm to determine what secondary

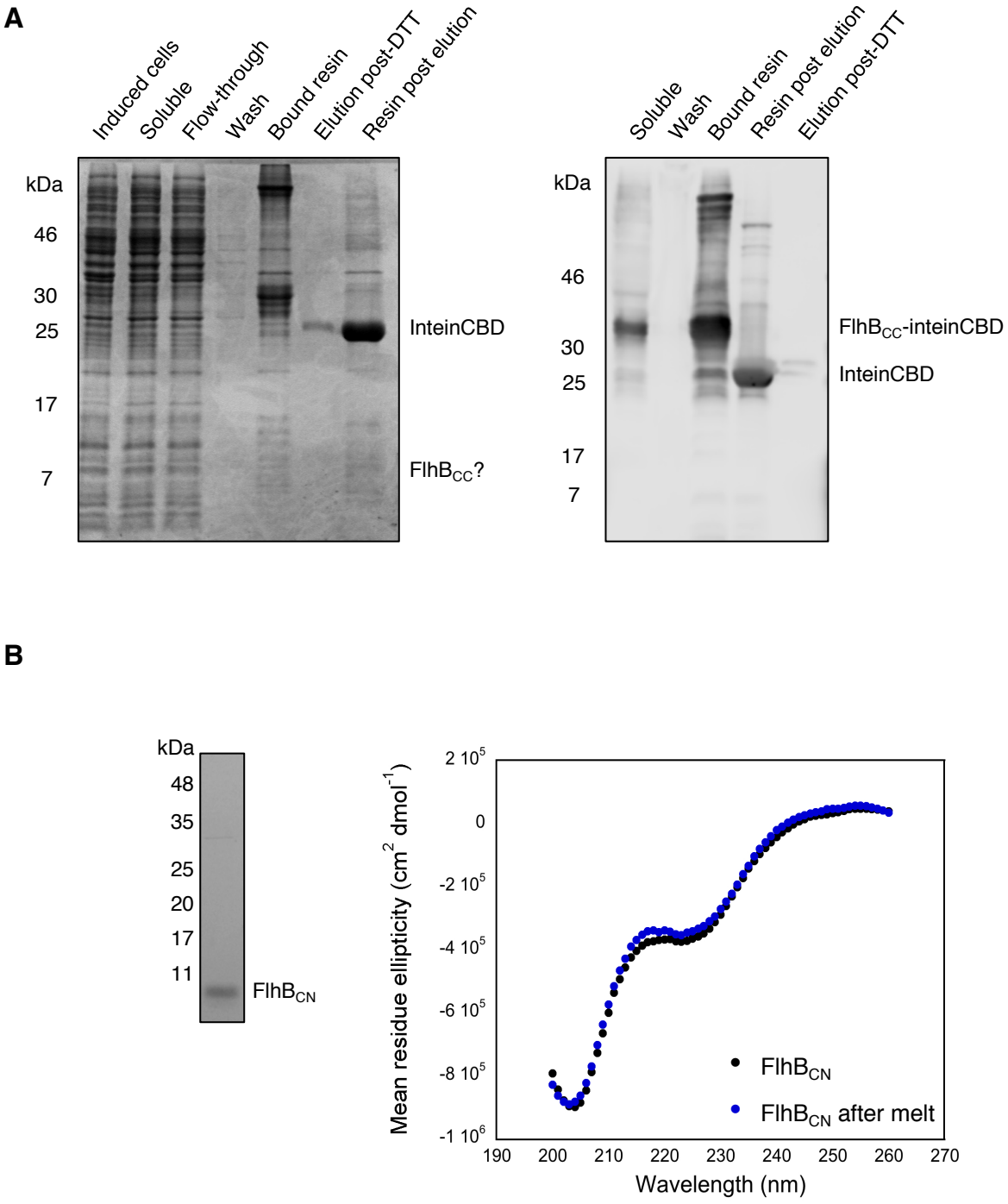


Figure 4.1 Independent expression and purification of FlhB cytosolic domains

(A) FlhB$_{CC}$-inteinCBD expressed in *E. coli* C41 and purified on chitin resin. **Left panel** is a Coomassie R250 stained 15% polyacrylamide gel, **right panel** is an immunoblot using mouse anti-chitin binding domain (anti-CBD) (New England Biolabs) to detect FlhB$_{CC}$-inteinCBD fusion protein and remaining inteinCBD. *Bound resin* is chitin resin bound with the FlhB$_{CC}$-inteinCBD fusion protein, *Elution post-DTT* is the eluate after DTT cleavage, and *Resin post elution* is protein bound to the resin after elution of the cleaved product. Samples were imaged using a Li-Cor Odyssey CLx and Image Studio Lite (Li-Cor).

(B) Purified FlhB$_{CN}$ was dialysed at 4°C overnight into sodium phosphate ($NaPO_4$), pH 7.4 buffer with 100 mM NaCl for use with circular dichroism (CD). Pure FlhB$_{CN}$ after dialysis is shown stained with Instant Blue Coomassie stain. CD scanning data from 260 nm to 200 nm at 25°C of FlhB$_{CN}$ showing the mean residue ellipticity before and after a thermal melt at 95°C.

structure, if any, it possessed in isolation from $FlhB_{CC}$. $FlhB_{CN}$ was subsequently incubated at 95°C before being returned to 25°C for a second scan from 260 nm to 200 nm to test $FlhB_{CN}$'s ability to refold after heat denaturation (Figure 4.1B).

Thermal melt and subsequent CD scan strongly suggest that $FlhB_{CN}$ does not adopt its native predominantly α-helical structure on its own in solution. The data is unclear as to what structure $FlhB_{CN}$ may adopt in solution, but it suggests that $FlhB_{CN}$ forms soluble, super-stable aggregates impervious to heat denaturation. Because neither $FlhB_{CC}$ expressed appropriately nor did $FlhB_{CN}$ adopt its native structure, further experiments investigating both proteins independently *in vitro* were not completed.

4.3 The autonomous switcher mutants surround the $FlhB_C$ β-sheet

Williams and colleagues identified a set of suppressor point mutations in FlhB that could overcome the effective loss of FliK to trigger the export specificity switch (243). These suppressor mutations, dubbed the autonomous switchers, are all located on $FlhB_{CC}$ (Figure 4.2A). Many of the suppressor mutations are found in the hydrophobic core of $FlhB_C$ surrounding the β-sheet, which has been proposed to contribute to holding the two $FlhB_C$ domains together post-autocleavage (Figure 4.2B) (241). The β-sheet comprises four β-strands – one from $FlhB_{CN}$ and three from $FlhB_{CC}$.

Some of the suppressor mutations introduce charged residues into the core of the protein or introduce larger hydrophobic residues into the core, potentially causing a rearrangement of the core. Others abolish non-covalent interactions, like hydrogen bonds, that could be important to holding the two polypeptides together post-autocleavage at the fT3SS. Finally, some of these mutations may alter the structure of $FlhB_{CC}$, resulting in different and perhaps weaker interactions between $FlhB_{CC}$ and $FlhB_{CN}$. I sought to investigate the role of these mutations in auto-cleaved $FlhB_C$ structural stability.

Williams and colleagues found little effect of the FlhB autonomous switcher mutations on cellular motility in the presence of FliK. I set out to identify mutations which, when introduced together, may also cause the greatest structural instability in $FlhB_C$. I hypothesised these joint mutations would result in greater defects of cellular motility

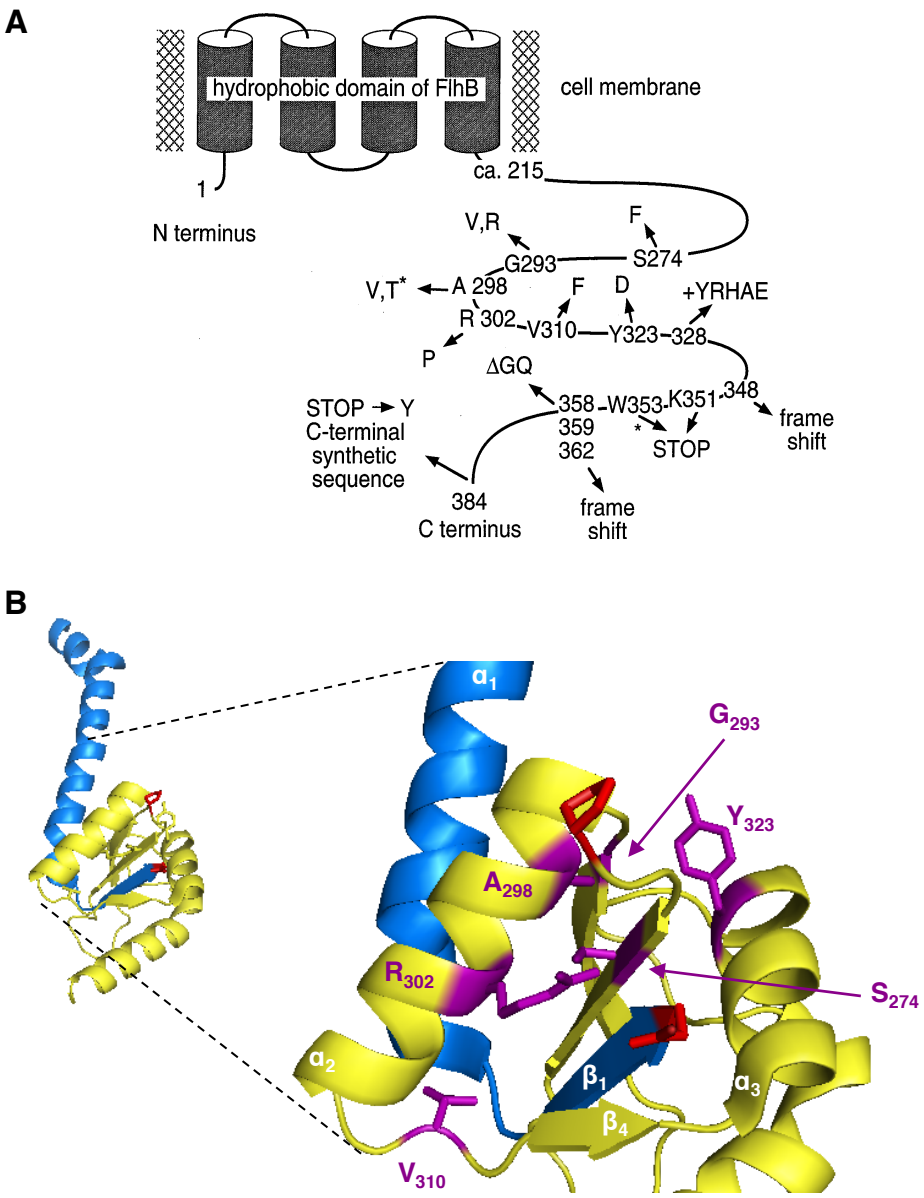


Figure 4.2 Suppressor mutations in FlhB, dubbed the "autonomous switchers", that can overcome the loss of FliK

(A) Autonomous switcher mutants are only found in $FlhB_{CC}$. Figure adapted from Williams et al., 1996.

(B) Autonomous switcher mutants mapped onto the atomic resolution structure of *Salmonella* $FlhB_{CC}$. The K_{351}Stop and W_{353}Stop mutants are not mapped to the structure. Tryptophan 353 represents the end of the folded structure as determined by X-ray crystallography, so these nonsense mutations are not shown. Insertion and frame shift mutations are likewise not shown.

in the presence of FliK. I also hypothesised that these joint mutations could be used to model the export specificity switch.

Serine$_{274}$ (S_{274}) and Arginine$_{302}$ (R_{302}) are mutated into phenylalanine and proline, respectively, in the mutants identified by Williams and colleagues (Figure 4.3). S_{274} lies downstream of the autocleavage site on β-strand$_2$ (Figure 4.3A). R_{302} sits in the center of α-helix$_2$ (Figure 4.3B), which a proline residue could disrupt. Mutations of either of these residues may alter autocleavage efficiency, therefore affecting the timing of the subunit specificity switch, or post-autocleavage structure, possibly resulting in a loss of FliK-mediated switching. Additionally, the side chains of the polar serine residue and charged arginine are within 4 Å of one another in the native FlhB$_C$ structure (Figure 4.3C), suggesting they may form a weak hydrogen bond that may be part of core stability. The autonomous switcher mutations abolish this hydrogen bond.

4.4 Two autonomous switcher mutations are better than one, but not as good as wild type

To be able to detect FlhB$_{CC}$ in the cell, my colleague, Owain Bryant, engineered a strain of *Salmonella enterica* serovar Typhimurium SJW1103 with a C-terminal 3xFLAG tag on FlhB$_{CC}$ expressed under native chromosomal control (FlhB$_{CT3xFLAG}$). I then introduced five sets of mutations of FlhB$_C$ into this strain using double lambda red recombinase-mediated recombination to engineer additional strains as described in section 2.7.1 (Figure 4.4A). A non-autocleavable variant of FlhB, FlhB $N_{269}A$, was engineered to serve as a negative control for *in vivo* experiments.

The effect on the cell's ability to swim of the FlhB autonomous switcher mutations in the presence of FliK was measured. The addition of a 3xFLAG tag to FlhB had little effect on swimming motility as cells expressing FlhB$_{CT3xFLAG}$ showed equivalent swimming motility to SJW1103 (Figure 4.4B). Cells expressing the autonomous switcher mutations individually had a slight swimming motility defect, but these cells swam comparatively to wild type (approximately 90% wild type swimming motility). Cells expressing the double autonomous switcher mutant ($S_{274}F$, $R_{302}P$) had a greater effect on motility, only about 60% as motile as wild type cells (Figure 4.4B).

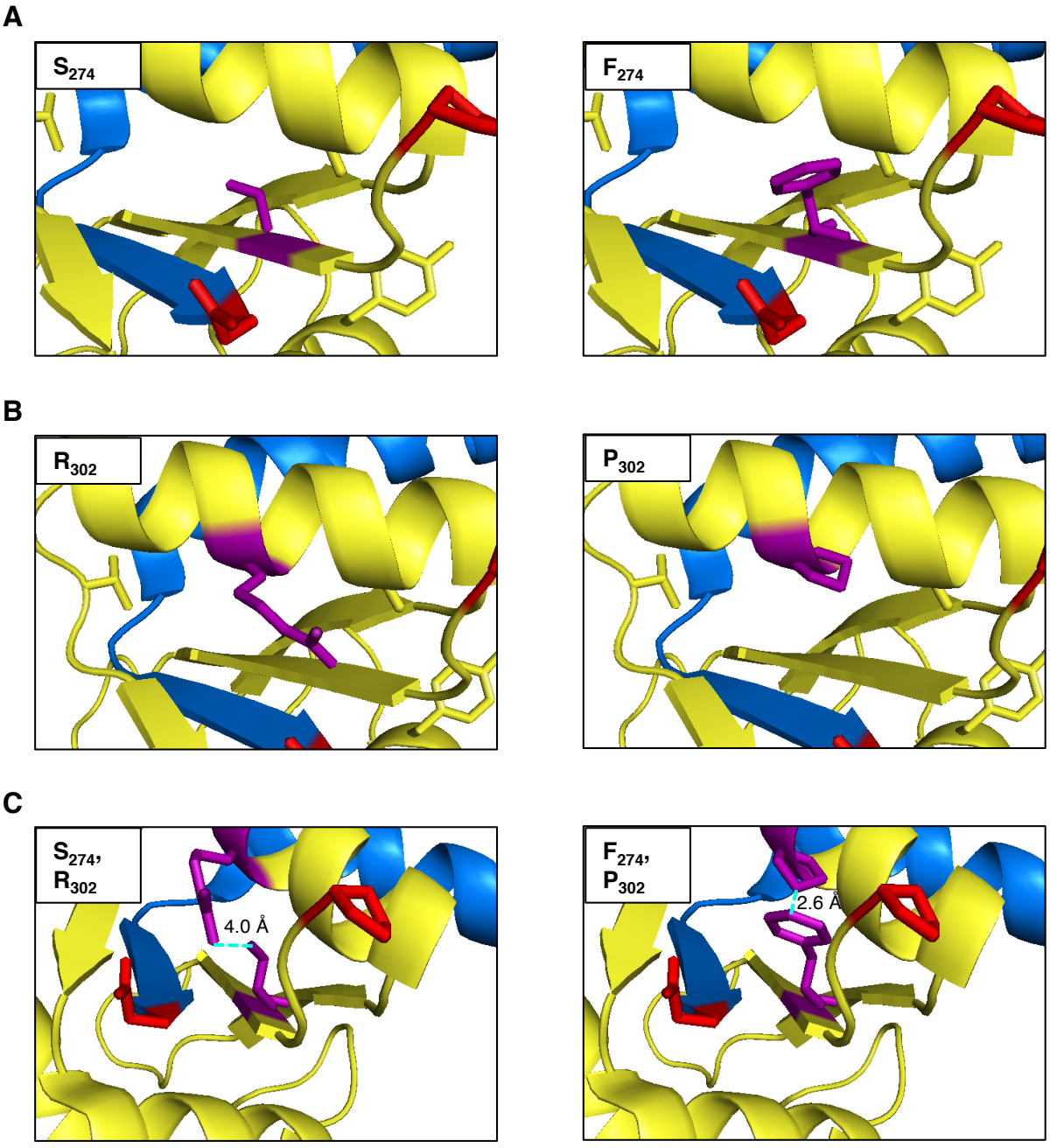


Figure 4.3 The autonomous switcher mutants selected for further study

(A) Autonomous switcher mutant $S_{274}F$ mapped onto the atomic resolution structure of *Salmonella* $FlhB_{CC}$. S_{274} sits on the β-strand immediately downstream of the autocleavage site. Its polar side chain sits in the core of the protein. The residues in red are the autocleavage site between N_{269} and P_{270}.

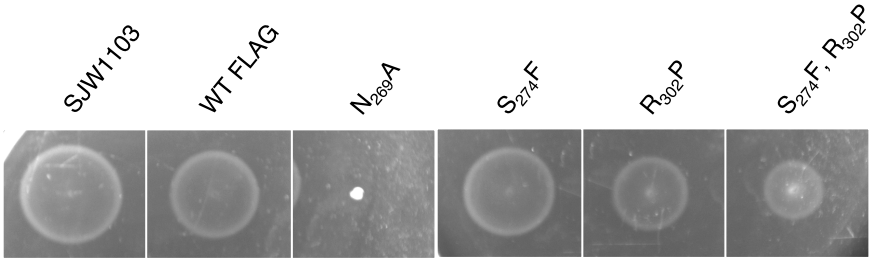
(B) Autonomous switcher mutant $R_{302}P$ mapped onto the atomic resolution structure of *Salmonella* $FlhB_{CC}$. R_{302} sits in the middle of α-$helix_2$, facing both the core and the surface of the protein, and above the autocleavage site. The residues in red are the autocleavage site between N_{269} and P_{270}.

(C) Autonomous switcher mutant $S_{274}F$, $R_{302}P$ mapped onto the atomic resolution structure of *Salmonella* $FlhB_{CC}$. The double replacement mutation moves these two residues 1.4 Å closer to one another.

A

Strain	Characteristic
SJW1103	*flhB*
WT FLAG	$flhB_{CT3xFLAG}$
$N_{269}A$	$flhB_{CT3xFLAG}$ $N_{269}A$, non-cleavable variant
PANA	$flhB_{CT3xFLAG}$ $P_{238}A$ $N_{269}A$, non-cleavable variant
$S_{274}F$	$flhB_{CT3xFLAG}$ $S_{274}F$
$R_{302}P$	$flhB_{CT3xFLAG}$ $R_{302}P$
$S_{274}F$, $R_{302}P$	$flhB_{CT3xFLAG}$ $S_{274}F$, $R_{302}P$

B



C

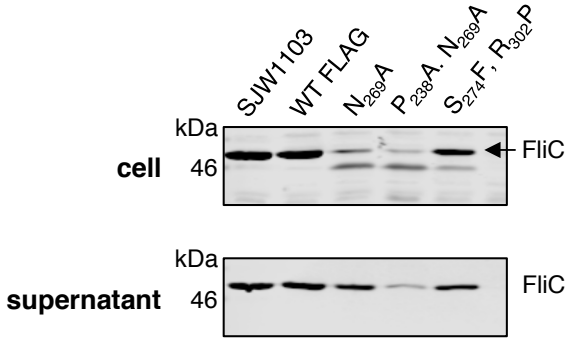


Figure 4.4 Autonomous switcher mutations have a slight effect on the motility phenotype and flagellin export

(A) List of SJW1103 strains used in motility and export assays.

(B) Motility assays of SJW1103$flhb_{Ct3xFLAG}$ variants. Strains were inoculated on swimming motility agar (0.25% w/v agar) and were incubated for 4 hours at 37°C.

(C) Immunoblot of export assay measuring FliC expression (cell) and export (supernatant) of SJW1103$flhb_{Ct3xFLAG}$ variants at A_{600} 1.0. The cell sample was collected from 1 ml of culture. The supernatant sample was filtered, cell-free supernatant from 1 ml of culture that was TCA precipitated to collect exported protein. Samples were diluted 1:100 and 1:50 in SDS buffer for cell and supernatant samples, respectively. All samples were standardised to A_{600}. FliC was detected using a 1:5000 dilution of Difco *Salmonella* H antiserum i (Becton, Dickinson and Company). Samples were imaged using a Li-Cor Odyssey CLx and Image Studio Lite (Li-Cor).

How might the double autonomous switcher mutant affect the cell's ability to swim? These mutations, which are known to affect FlhB's ability to autocleavable, may impair the expression and export of flagellin (FliC) by mistiming the export specificity switch. Cells expressing wild type FlhB and $FlhB_{CT3xFLAG}$ expressed and exported similar levels of FliC, as expected based on their ability to swim (Figure 4.4C). Cells carrying the double $S_{274}F$, $R_{302}P$ mutant expressed slightly lower levels of FliC, which resulted in decreased export, although the effect is not great (Figure 4.4C). This small decrease in flagellin expression and export may explain the defect in motility, but the data suggest the cell can recover.

While cells expressing the non-autocleavable *flhB* $N_{269}A$ variant could not swim, they expressed low levels of *fliC* and were able to export FliC into the supernatant. This is surprising as past reports have shown the *flhB* $N_{269}A$ mutant incapable of exporting FliC (193). FlhB does possess a known secondary autocleavage site between D_{237} and P_{238}, although this site is not known to be biologically relevant (193). P_{238} was mutated along with N_{269}, resulting in the "PANA" strain. This *flhB* PANA mutant is totally autocleavage deficient and should not switch from early to late subunit export. FliC expression and export was greatly reduced in the PANA strain (Figure 4.4C).

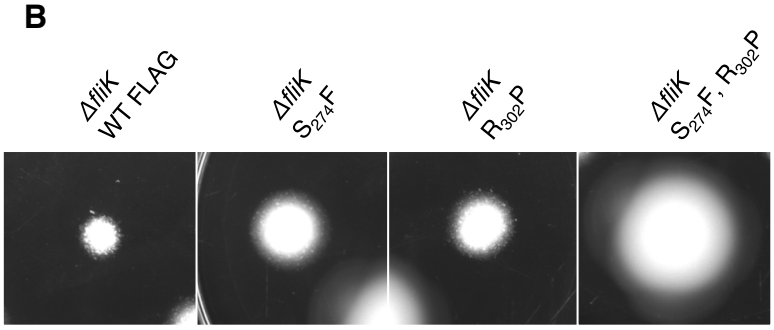
The double autonomous switcher mutant showed minor phenotypic effect in the presence of FliK. To test its effect on swimming motility and flagellin export without FliK, FliK was effectively deleted by replacement of the *fliK* gene with a $kanamycin^R$ cassette (Figure 4.5A). In past work, the autonomous switcher mutants were evaluated in strains expressing a C-terminal truncation of FliK (residues 1-300), not enough to cause the export specificity switch but still enough for FliK secretion (243). Here, the full *fliK* gene was replaced with the $kanamycin^R$ cassette, leaving only the extreme 5' and 3' gene sequence to allow for native operon control.

Cells expressing the individual autonomous switcher mutants swam poorly in the absence of FliK (Figure 4.5B). Cells expressing the double autonomous switcher mutant showed greater swimming motility (Figure 4.5B); however, they did not swim near as well as wild type *Salmonella* with FliK. Cells expressing the individual and double autonomous switcher mutants all took 20 hours before motility could be

A

Strain	Characteristic
ΔfliK WT FLAG	*ΔfliK::Kan^R* *flhB$_{CT3xFLAG}$*
ΔfliK $S_{274}F$	*ΔfliK::Kan^R* *flhB$_{CT3xFLAG}$* $S_{274}F$
ΔfliK $R_{302}P$	*ΔfliK::Kan^R* *flhB$_{CT3xFLAG}$* $R_{302}P$
ΔfliK $S_{274}F$, $R_{302}P$	*ΔfliK::Kan^R* *flhB$_{CT3xFLAG}$* $S_{274}F$, $R_{302}P$

B



C

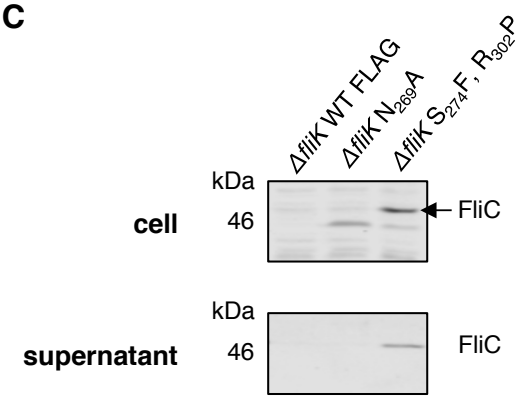


Figure 4.5 Two autonomous switcher mutations are better than one

(A) List of SJW1103 strains used in motility and export assays.

(B) Motility assays of SJW1103*ΔfliK::Kan^R flhb$_{Ct3xFLAG}$* variants. Strains were inoculated on swimming motility agar (0.25% w/v agar) and were incubated for 20 hours at 37°C.

(C) Immunoblot of export assay measuring FliC expression (cell) and export (supernatant) of SJW1103*ΔfliK::Kan^R flhb$_{Ct3xFLAG}$* variants at A_{600} 1.0. The cell sample was collected from 1 ml of culture. The supernatant sample was filtered, cell-free supernatant from 1 ml of culture that was TCA precipitated to collect exported protein. Samples were diluted 1:100 and 1:50 in SDS buffer for cell and supernatant samples, respectively. All samples were standardised to A_{600}. FliC was detected using a 1:5000 dilution of Difco *Salmonella* H antiserum i (Becton, Dickinson and Company). Samples were imaged using a Li-Cor Odyssey CLx and Image Studio Lite (Li-Cor).

observed. At this stage, wild type *Salmonella* cells would have overtaken the entire motility plate and, therefore, was not included in these long-term experiments. As expected, cells expressing wild type *flhB* were unable to overcome the loss of FliK, displaying some outward growth after 20 hours but no signs of swimming motility (Figure 4.5B).

Cells expressing the *flhB* double autonomous switcher mutant in the absence of FliK showed low levels of FliC expression and export (Figure 4.5C). The assay evaluated *fliC* expression at A_{600} 1.0, or about three hours of growth. As it takes cells expressing the double autonomous switcher mutant around 20 hours to swim in the absence of FliK, it is not surprising that FliC expression and export is weak at an early stage of growth. Further work is needed to evaluate FliC expression and export in the double autonomous switcher mutant after 20 hours of growth.

Based on swimming motility and FliC export assays, it is clear the FlhB double autonomous switcher mutant has a greater effect on the cell than individual autonomous switcher mutants. How might these mutations impact flagellar subunit export and assembly? Might they mistime the switch by disrupting the $FlhB_C$ structure?

4.5 Autonomous switcher mutants are less stable in vitro

Four autonomous switcher mutations were cloned into the pTXBI expression vector to study the effects of the mutations on $FlhB_C$ stability *in vitro*. The four mutations were as follows – $S_{274}F$, $R_{302}P$, $S_{274}F$ $R_{302}P$ combined, and $W_{353}T$. $W_{353}T$ was constructed to mimic the W_{353}Stop mutation that excludes the last 30 residues of $FlhB_C$ (residues 354-383). Wild type FlhB and non-autocleavable FlhB $N_{269}A$ were also included in the stability study. All variants of $FlhB_C$ expressed well and could be purified using the inteinCBD system (Figure 4.6A).

CD experiments were planned and conducted with Dr. Katie Kemplen. Dr. Kemplen analysed the data using KaleidaGraph to generate graphs and fit the equation provided in the Materials and Methods to the CD thermal melt data. Wild type $FlhB_C$ was initially scanned using CD from 260 nm to 200 nm at 25°C and 35°C. The protein showed a slightly greater signal at 35°C, not unexpected as *Salmonella*'s preferred

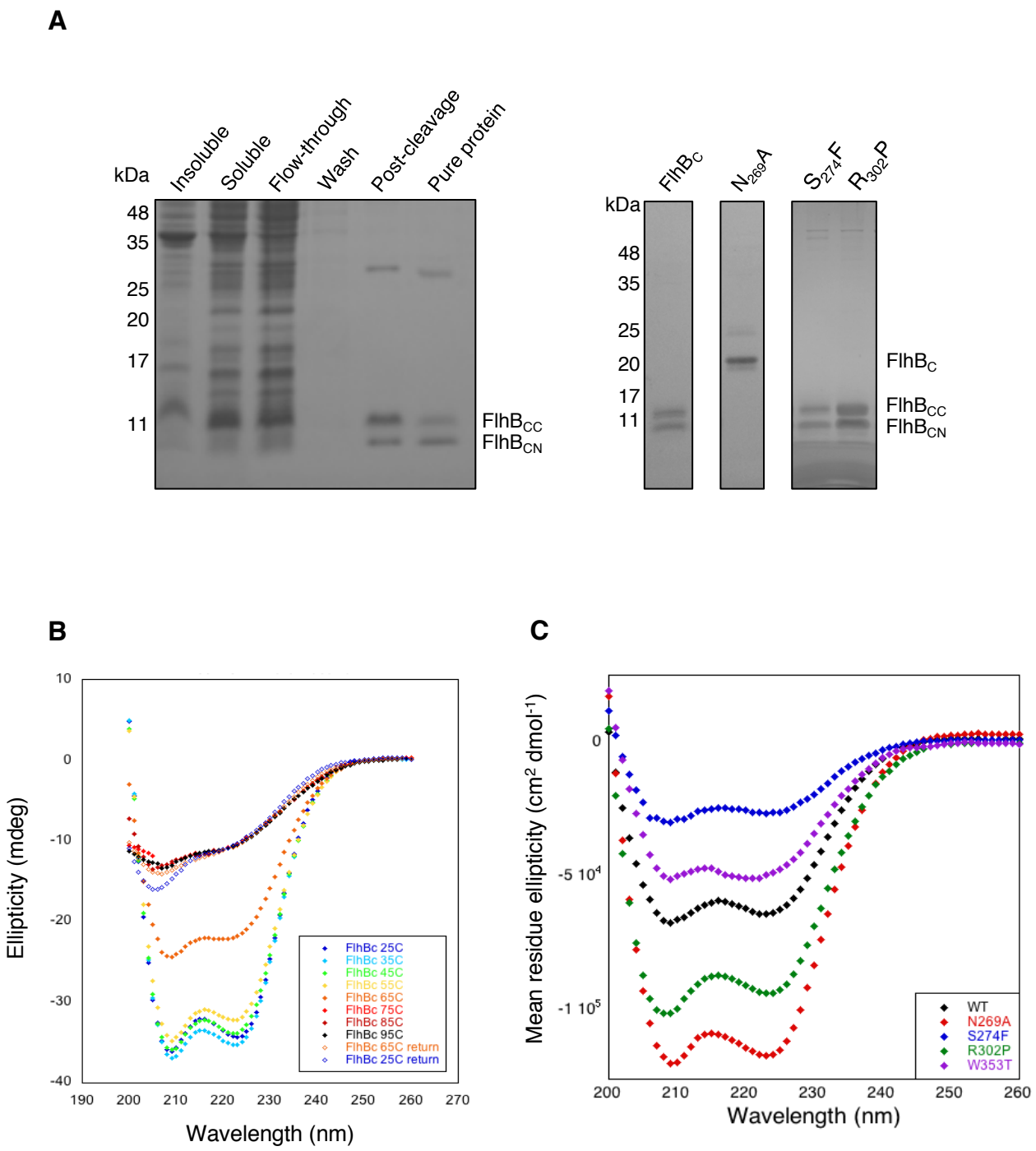


Figure 4.6 Autonomous switchers mutants do not have substantial structural differences from wild type $FlhB_C$

(A) $FlhB_C$ $W_{353}T$ tagged with inteinCBD purified on chitin resin following method described in materials and methods. Samples were run on 15% SDS-PAGE and stained with Instant Blue coomassie stain. Purified protein was dialysed at 4°C overnight into 20 mM sodium phosphate ($NaPO_4$), pH 7.4 buffer with 100 mM NaCl for use with CD. Post-cleavage represents $FlhB_C$ $W_{353}T$ after cleavage from the inteinCBD. Pure protein is $FlhB_C$ $W_{353}T$ after dialysis. Panels on the right are the pure protein inputs used in CD. All $FlhB_C$ variants were purified and dialysed in the same manner.

(B) Scanning thermal melts of wild type $FlhB_C$ from 260 nm to 200 nm. Temperature was increased by 10°C and allowed to equilibrate for 10 minutes before scanning $FlhB_C$ again. $FlhB_C$ was completely unfolded at 75°C and did not refold as temperature was lowered.

(C) CD scanning data from 260 nm to 200 nm of $FlhB_C$ variants. Autocleaved $FlhB_C$ (WT), non-autocleaved $FlhB_C$ ($N_{269}A$), $FlhB_C$ $S_{274}F$, and $FlhB_C$ $W_{353}T$ show similar structure based on scans. $FlhB_C$ $R_{302}P$ does show a shift in the curve around 208 nm, suggesting a different structure to other FlhB variants.

growth temperature is 37°C. Wild type $FlhB_C$ was further scanned in intervals of 10°C, up to 95°C, to unfold the protein. After the melt at 95°C, $FlhB_C$ was incubated at 65°C and then 25°C to investigate if it could refold as it was returned to room temperature. The CD signal for wild type $FlhB_C$ began to weaken at 55°C and reached its lowest signal around 75°C, suggesting the protein is fully denatured around this temperature. The CD signal never recovered as the temperature was lowered, indicating that $FlhB_C$ cannot refold on its own once unfolded by heat denaturation (Figure 4.6B).

Each of the FlhB variants were scanned from 260 nm to 200 nm at 25°C to determine if they were structurally similar. Most of the variants returned a similar curve as wild type, autocleavable $FlhB_C$ (Figure 4.6C). This indicates that some of the individual mutations do not have a major effect on $FlhB_C$ structure. One autonomous switcher mutant, $R_{302}P$, showed a shift around 208 nm in the CD curve. Although it is not possible from the CD data recorded to conclude what structural differences exist between wild type $FlhB_C$ and $FlhB_C$ $R_{302}P$, this shift suggests there is some structural change between these two variants. Based on the signal and the location of the point mutation, it is possible that the observed difference is due to a change in α-$helix_2$.

Unfortunately, the double autonomous switcher mutant, $S_{274}F$ $R_{302}P$, precipitated out of solution when dialysed against the CD experimental buffer (20mM sodium phosphate, pH 7.4, 100 mM NaCl) at 4°C (Figure 4.7). Because this protein was unstable in buffer suitable for circular dichroism (CD), it could not be studied further in these experiments.

The *in vitro* stability of each FlhB variant was determined by CD at 222 nm and measuring the melting temperature of the protein by increasing the experimental temperature from 25°C to 95°C in 1°C intervals. Each variant was also studied for the ability to refold by decreasing temperature from 95°C to 25°C in 1°C intervals after an incubation step at 95°C. Wild type and non-autocleavable $FlhB_C$ had similar melting temperatures near 80°C (Figure 4.8A and 4.8B). Non-autocleavable $FlhB_C$ was slightly more thermostable, which is not surprising as the protein possesses an additional (covalent) peptide bond.

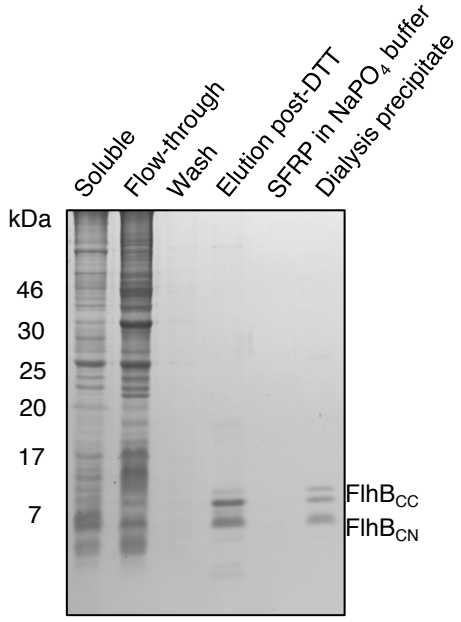


Figure 4.7 The double autonomous switcher mutant was not stable in buffer suitable for CD

$FlhB_C$ $S_{274}F$, $R_{302}P$ tagged with inteinCBD purified on chitin resin. Samples were run on 15% SDS-PAGE and stained with Instant Blue coomassie stain. After purification in 40 mM Tris, pH 7.4 buffer with 1 M NaCl, protein was dialysed at 4°C overnight into 20 mM sodium phosphate ($NaPO_4$), pH 7.4 buffer with 100 mM NaCl for use with CD. The double autonomous switcher mutant was unstable in the low-salt sodium phosphate buffer used for CD and precipitated out of solution. Therefore, this mutant could not be analysed further by CD.

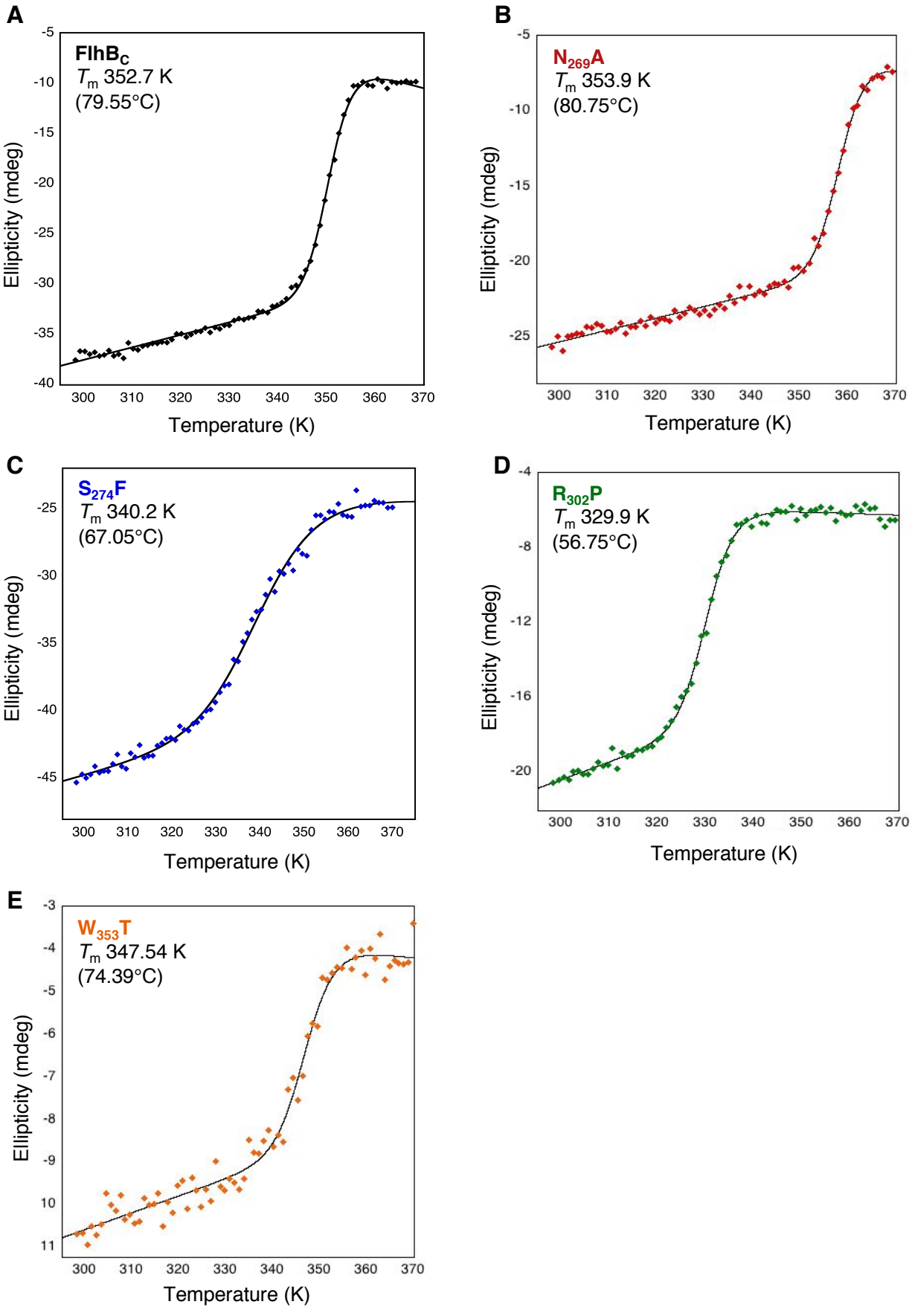


Figure 4.8 $FlhB_C$ autonomous switcher mutants alter its thermal stability

(A-E) Thermal melts of $FlhB_C$ variants from 298.15 K (25°C) to 368.15 K (95°C) measured at 222 nm by CD. Proteins were subsequently cooled to 298.15 K (25°C), but no refolding was observed as previously shown for wild type $FlhB_C$ (Figure 4.6B).

Unlike the non-autocleavable FlhB variant, a single autonomous switcher mutation was enough to substantially decrease the melting temperature of FlhB. $S_{274}F$ was 12°C less stable (Figure 4.8C). $R_{302}P$ was 23°C less stable than wild type $FlhB_C$ (Figure 4.5D). The $W_{353}T$ mutant was about 5°C less stable than wild type $FlhB_C$ (Figure 4.8E). The loss of thermal stability among the autonomous switcher mutants indicates that these mutations are disrupting the $FlhB_C$ polypeptides, even if they are not radically changing the folded structure.

Structural changes caused by these autonomous switcher mutations were studied by attempting to crystallise $FlhB_C$ $R_{302}P$ using the crystal screens in section 2.20. $FlhB_C$ $R_{302}P$ was chosen as the CD data suggested it may be structurally different from wild type FlhB and it had the most severe effect on $FlhB_C$ melting temperature. $FlhB_C$ $R_{302}P$ failed to crystallise in the three screens used to crystallise $FlhB_C$ successfully. Further investigation of the structural differences between $FlhB_C$ and the autonomous switchers is needed to study how these point mutations effect the structure of $FlhB_C$.

4.6 Autonomous switcher mutants are less stable after autocleavage in vivo

The autonomous switcher mutants proved to be less stable *in vitro* than autocleavable, wild type FlhB and non-autocleavable FlhB. Do these mutations have a similar effect on FlhB stability in the cell? To observe FlhB stability *in vivo*, *Salmonella ΔfliK::Kan*^{R} cells expressing 3xFLAG-tagged wild type FlhB, 3xFLAG-tagged non-autocleavable FlhB $N_{269}A$, and a 3xFLAG-tagged FlhB double autonomous switcher mutant were grown to late log and stationary phase as measured by A_{600}. Wild type $FlhB_C$ has completely autocleaved by late log stage as detected by the single $FlhB_{CC\ CT3xFLAG}$ band and a lack of a corresponding full length $FlhB_{Ct3xFLAG}$ (uncleaved) band (Figure 4.9). This $FlhB_{CC\ CT3xFLAG}$ band is observed in both late log and stationary phases. The non-autocleavable $FlhB_C$, by contrast, has not autocleaved, as is expected, since only a full-length band for $FlhB_{CT3xFLAG}$ was observed (Figure 4.9).

The double autonomous switcher mutant shows inefficient autocleavage *in vivo* by late log stage, but is fully autocleaved by stationary phase. A band for $FlhB_{CC\ Ct3xFLAG}$ cannot be seen either in late log phase or in stationary phase for the double autonomous switcher mutant (Figure 4.9). This suggests that $FlhB_{CC}$ is unstable after

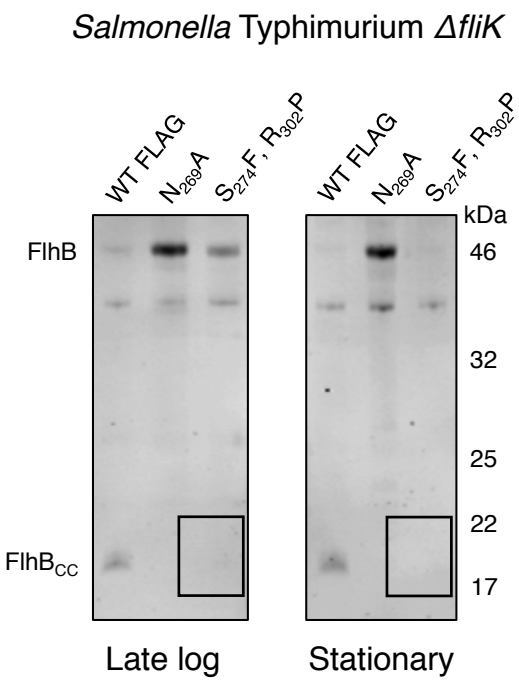


Figure 4.9 *In vivo* stability of FlhB in late log and stationary phase

The box shows where the $FlhB_{CC}$ band should appear. *Salmonella* Typhimurium *ΔfliK::Kan*^{R} cells with C-terminally FLAG tagged *flhB* variants were grown in LB broth with kanamycin 50 μg ml^{-1} until an A_{600} of 1.0 for late log phase and A_{600} 1.5-2.0 for stationary phase. Whole cell samples were collected and standardised to A_{600}. Samples were analysed by immunoblotting using mouse anti-FLAG M2 antibody (Sigma) and donkey anti-mouse infrared antibody (Li-Cor). Samples were imaged using a Li-Cor Odyssey CLx and Image Studio Lite (Li-Cor). $FlhB_{CN}$ was not detected with antibody in these experiments.

autocleavage as a result of the double autonomous switcher mutations and agrees with the previous *in vitro* findings that these mutations destabilise FlhB.

4.7 FliK leads to the loss of $FlhB_{CC}$ from the membrane

The autonomous switcher mutations destabilised $FlhB_C$ *in vitro* and *in vivo* in the absence of FliK. They hint that the role of FliK in the export specificity switch might be to destabilise the non-covalent interactions that bind $FlhB_{CC}$ to the rest of FlhB. To test this hypothesis, whole cell samples and small scale fractionation experiments at late log phase (A_{600} 1.0) were carried out with cells expressing 3xFLAG-tagged wild type and non-autocleavable $FlhB_C$ in the presence and absence of FliK. FlgN, a soluble cytoplasmic protein, was used as a control for the fractionation experiments.

Wild type $FlhB_C$ is completely autocleaved in late log phase whether FliK is present or not in whole cell samples. Thus, FliK doesn't seem to have much of an effect on FlhB autocleavage (Figure 4.10). In the absence of FliK, a distinct $FlhB_{CC\ Ct3xFLAG}$ band can be seen in the whole cell samples as previously noted (Figure 4.10). The presence of FliK leads to a 2.4-fold loss of the $FlhB_{CC\ Ct3xFLAG}$ band (Figure 4.10). This result is statistically significant (p-value= 0.006) across four biological replicates.

Fractionation experiments in the presence of FliK were used to determine what might happen to $FlhB_{CC}$ after it dissociates. The $FlhB_{CC}$ band does not appear in the periplasmic fraction, which is not surprising (Figure 4.11B). The $FlhB_{CC}$ band does appear in the cytosol fraction, but so does non-autocleaved FlhB ($N_{269}A$), suggesting the cytosol fraction was slightly contaminated with insoluble and membrane protein (Figure 4.11C). The insoluble fraction, which contains membrane and insoluble cytoplasmic protein, mimics what is seen in the whole cell fraction, where $FlhB_{CC}$ is decreased 2.4-fold in the presence of FliK (Figure 4.11A and Figure 4.11D). FlhB is a transmembrane protein, so it is reasonable to suggest that $FlhB_{CC}$ is lost from the membrane before being degraded by proteases. These data support that hypothesis. Alternatively, $FlhB_{CC}$ might be too unstable to be detected after dissociation from FlhB.

The separation of fractions was not as clean as desired as all fractions contained some FlgN, although it was mostly detected in the cytosol fraction as expected. Very little

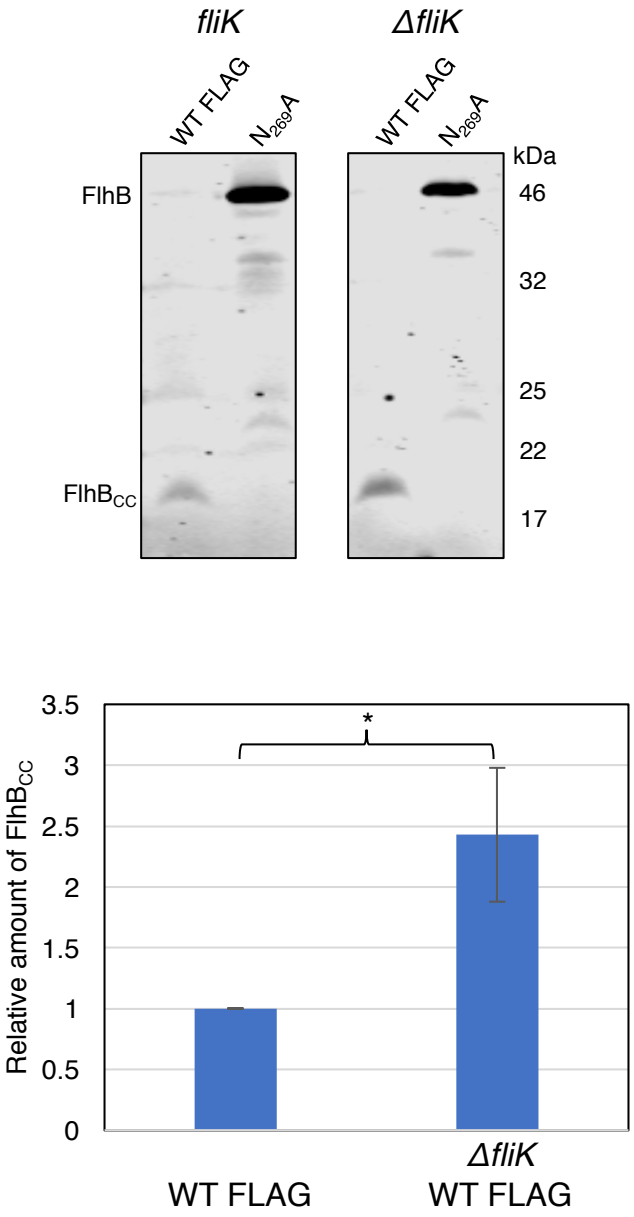


Figure 4.10 $FlhB_{CC}$ dissociates from FlhB in a FliK-dependent manner

2.4 fold less $FlhB_{CC}$ is observed when FliK is present. *Salmonella* Typhimurium wild type and *ΔfliK::Kan*^{R} cells with C-terminally FLAG tagged *flhB* variants were grown in LB broth until an A_{600} of 1.0. 1 ml whole cell samples were collected and standardised to A_{600}. Samples were analysed by immunoblotting using mouse anti-FLAG M2 antibody (Sigma) and donkey anti-mouse infrared antibody (Li-Cor). $FlhB_{CN}$ was not detected with antibody in these experiments. The relative amount of $FlhB_{CC}$ with standard error of the mean was calculated from four biological replicates and analysed using a Li-Cor Odyssey CLx and Image Studio Lite (Li-Cor). The relative amount of $FlhB_{CC}$ in WT FLAG samples was always counted as 1.0. * indicates that the p-value < 0.05.

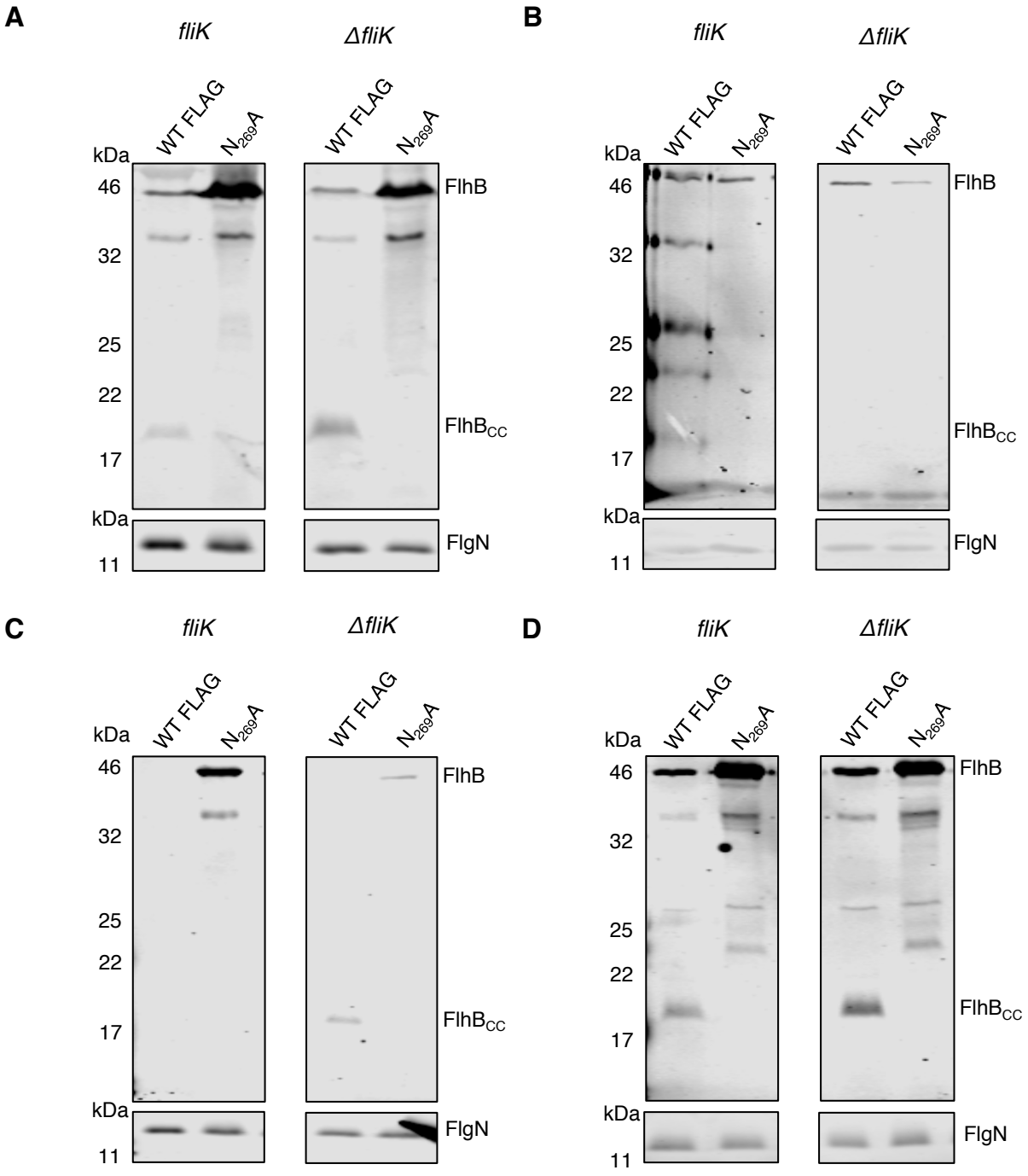


Figure 4.11 $FlhB_{CC}$ disappears from the membrane

Salmonella Typhimurium wild type and *ΔfliK::Kan*^{R} cells with C-terminally FLAG tagged *flhB* variants were grown in 10 ml LB broth until an A_{600} of 1.0. 1 ml whole cell samples were collected and fractionated using osmotic shock. Samples were Samples were analysed by immunoblotting using mouse anti-FLAG M2 antibody (Sigma) and rabbit anti-FlgN. Donkey anti-mouse and donkey anti-rabbit infrared antibody (Li-Cor) were used as secondary antibodies. Samples were imaged using a Li-Cor Odyssey CLx and Image Studio Lite (Li-Cor). Panels beneath each blot show the amount of FlgN, a soluble, cytoplasmic flagellar subunit, in each sample.

(A) Whole cell

(B) Periplasmic fraction

(C) Cytosol fraction

(D) Insoluble fraction

FlgN appeared in the periplasmic fraction, confirming that osmotic shock used to separate periplasm contents from the cell was successful overall. The presence of FlgN in the insoluble fraction could indicate that it started to aggregate, but is more probably a result of the slow centrifuge speed (17,000 xg) used to separate the fractions.

After dissociation, $FlhB_{CC}$ does not appear elsewhere in the cell as assayed by fractionation. Past reports suggest that Yersinia $SctU_{CC}$ contains a possible secretion signal and is secreted after the export specificity switch is complete (198, 199). A secretion assay using 100 ml of culture was conducted to see if $FlhB_{CC}$ was secreted after being removed from the membrane. Wild type *Salmonella* expressing FlhB and FliK was grown to an A_{600} of 1.0. Although wild type FlhB did not autocleave fully by this stage, a $FlhB_{CC}$ band was observed in the whole cell. No corresponding $FlhB_{CC}$ band was detected in the supernatant, suggesting that $FlhB_{CC}$ is not secreted after its removal from the membrane (Figure 4.12).

4.8 How might FliK lead to the loss of $FlhB_{CC}$ from FlhB

The data suggest that $FlhB_{CC}$ dissociates from the rest of FlhB in a FliK-dependent manner and that the autonomous switcher mutations mimic the role of FliK in the export specificity switch. How might FliK cause the dissociation between $FlhB_{CC}$ and FlhB? Evidence for a direct interaction between the FliK T3S4 domain and $FlhB_{CC}$ is weak, so a few hypotheses were tested to determine how this FliK-dependent loss of $FlhB_{CC}$ may occur.

Firstly, full-length FliK was expressed in *E. coli* C41 to determine if FliK could cause $FlhB_{CC}$ to dissociate from FlhB *in vitro*. *E. coli* C41 cells were resuspended in buffer and lysed to collect the soluble fraction containing FliK. Pure $FlhB_{C\ Ct3xFLAG}$ (1 μM) was incubated with both C41 soluble lysate expressing empty pET15b vector and C41 soluble lysate containing full-length FliK. Samples were collected at different time points and analysed by immunoblotting to detect any loss of $FlhB_{CC}$. No loss of $FlhB_{CC}$ could be observed *in vitro* after 180 minutes, approximately the same length of time as the *in vivo* cell-based assays that originally showed the FliK-dependent loss of $FlhB_{CC}$ (Figure 4.13).

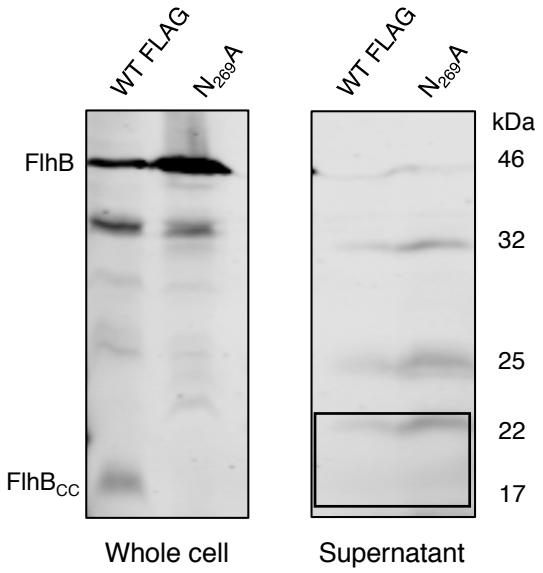


Figure 4.12 $FlhB_{CC}$ is not secreted after dissociation from FlhB

Salmonella Typhimurium wild type cells with C-terminally FLAG tagged *flhB* variants were grown in 100 ml LB broth until an A_{600} of 1.0. Cells of 1 ml broth culture were collected for whole cell samples. Supernatant samples were TCA precipitated from remaining 99 ml of culture. Samples were analysed by immunoblotting using mouse anti-FLAG M2 antibody (Sigma) and donkey anti-mouse infrared antibody (Li-Cor). Samples were imaged using a Li-Cor Odyssey CLx and Image Studio Lite (Li-Cor). $FlhB_{CN}$ was not detected with antibody in these experiments.

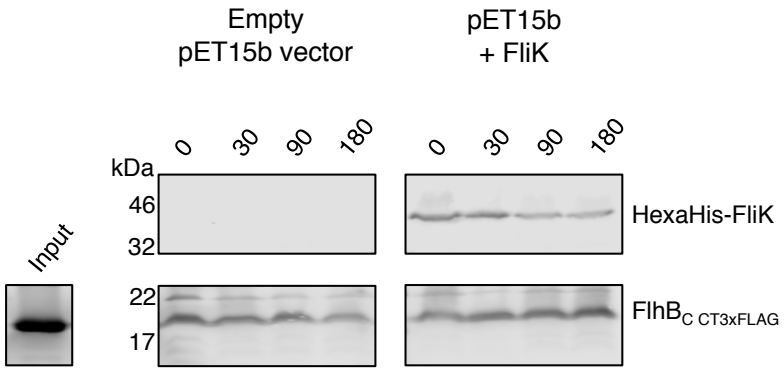


Figure 4.13 FliK does not cause the loss of $FlhB_{cc}$ *in vitro*

Immunoblot of *in vitro* assay of purified $FlhB_{C\ Ct3xFLAG}$ (input) incubated for 0, 30, 90, and 180 minutes at 25°C with C41 soluble lysate ± HexaHis-FliK. Left panel is the C41 empty vector soluble lysate control. $FlhB_{C\ Ct3xFLAG}$ was detected with mouse anti-FLAG M2 antibody. HexaHis-FliK was detected with mouse PentaHis antibody (Qiagen). Donkey anti-mouse infrared antibody (Li-Cor) was used to detect primary antibodies. Samples were imaged using a Li-Cor Odyssey CLx and Image Studio Lite (Li-Cor).

This *in vitro* experiment failed to demonstrate that the loss of $FlhB_{CC}$ was FliK-dependent, unlike the *in vivo* experiment that showed a 2.4-fold loss of $FlhB_{CC}$ in the presence of FliK. Minamino and colleagues used genetic analysis to conclude that FliK residues 300-350 and 401-405 are required for the export specificity switch, the so-called T3S4 domain (202). Residues 300-350 correspond to the central globular $FliK_C$ domain (203). Recombinant plasmid (pBAD18amp) carrying $FliK_{270-405}$ was constructed to test whether this T3S4 domain alone could interact with FlhB to cause the export specificity switch.

Salmonella Typhimurium *ΔfliK::FRT* cells expressing $fliK_{270}$ induced with 0.02% w/v L-arabinose did not swim at 37°C, while cells expressing wild type *fliK* could complement swimming motility in the same conditions (Figure 4.14A). That was unsurprising given previously reported results that $fliK_C$ was not enough to complement swimming motility (204).

Along with Dr. Kemplen, I planned and performed a pulldown experiment to test whether $FliK_{270}$ might interact with pure $FlhB_C$. These experiments were performed with HexaHis-tagged FliK and $FliK_{270}$ immobilised on nickel resin. Resin-bound FliK and $FliK_{270}$ were then incubated with purified, untagged $FlhB_C$. $FlhB_C$ bound to the empty nickel agarose resin, but bound more greatly to FliK and $FliK_{270}$ (Figure 4.14B). $FlhB_C$ bound to FliK and $FliK_{270}$ in equal amounts, suggesting $FliK_{270}$ is able to form a stable complex with $FlhB_C$. This is one of the first biochemical results to show a direct interaction between $FliK_C$ and $FlhB_C$. The high amount of background $FlhB_C$ binding to nickel-NTA resin, however, makes these results difficult to analyse conclusively. That, coupled with the lack of motility from the $FliK_{270}$ construct, made this a less attractive truncation to study. Attention was returned to full-length FliK.

The autocleavage of $FlhB_C$ leaves a surface-exposed proline residue at the start of $FlhB_{CC}$. We hypothesised that FliK may act as a proline aminopeptidase and remove this proline residue to begin to destabilise $FlhB_C$. L-proline p-nitroanilide trifluoroactetate salt mimics a free N-terminal proline residue. In the presence of a proline aminopeptidase, L-proline will be cleaved from the p-nitroanilide, resulting in a colourimetric change that can be measured at 405 nm.

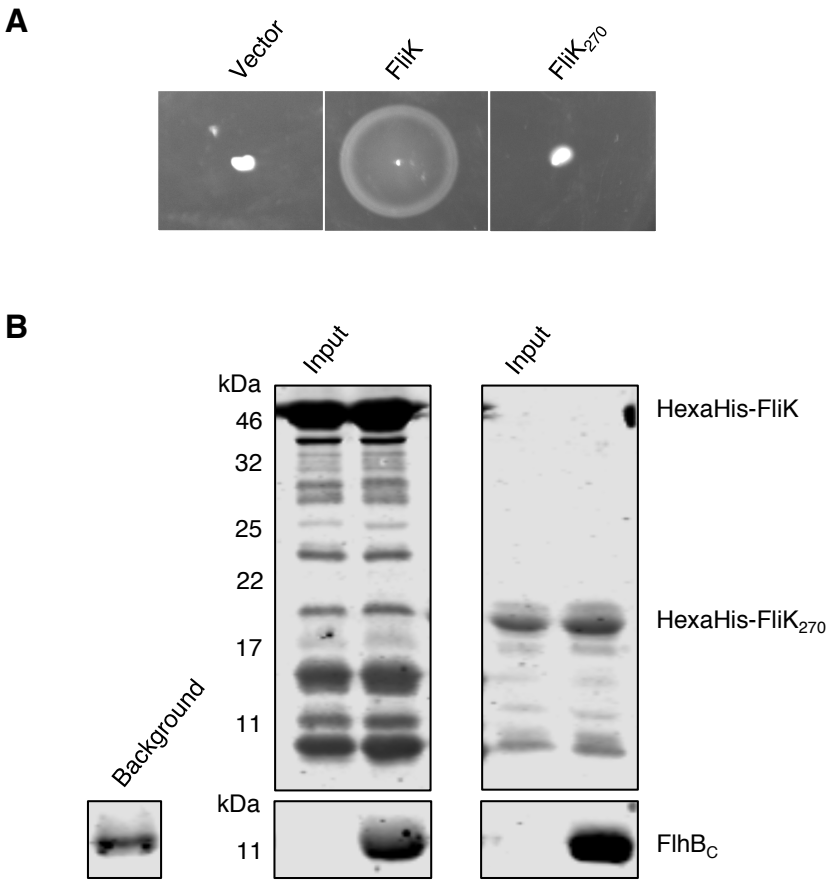


Figure 4.14 The FliK T3S4 domain alone can not cause the switch, but can bind to $FlhB_C$

(A) *$fliK_{270}$* is not able to recover motility in a SJW1103*ΔfliK::FRT* strain. pBAD18amp served as the empty vector control. Protein expression was induced with 0.02% w/v L-arabinose for four hours at 37°C.

(B) $FliK_{270}$ binds $FlhB_C$. The left most band, labeled “background”, is $FlhB_C$ binding to empty nickel-NTA resin. HexaHis-FliK and HexaHis-$FliK_{270}$ were purified on nickel-NTA resin (Qiagen) and washed with 10 CV. A sample of FliK bound on resin was taken to serve as the “input”. Pure $FlhB_C$ was then added to the remaining resin. The FliK-$FlhB_C$ complex was washed with 5 CV and proteins were boiled from the resin. Samples were analysed by immunoblotting using PentaHis antibody (Qiagen) and anti-$FlhB_C$ antibody. Samples were imaged using a Li-Cor Odyssey CLx and Image Studio Lite (Li-Cor).

Purified FliK was incubated with 1 mM L-proline p-nitroanilide trifluoroactetate at 20°C or 50°C for 30 minutes. C41 pET15b soluble lysate in PBS was used as the positive control as *E. coli* has native proline aminopeptidases. The reaction was stopped with the addition of acetic acid. Each sample was analysed at A_{405}. The C41 pET15b soluble lysate had the greatest aminopeptidase activity at 20°C. Incubating the solutions at 50°C resulted in about a twofold loss of enzymatic activity. After 30 minutes incubation, FliK showed no effect on L-proline p-nitroanilide trifluoroactetate at either 20°C or 50°C (Figure 4.15). Therefore, FliK is most likely not a proline aminopeptidase.

Finally, Mizuno et al 2011 proposed that the acidic loops of $FliK_C$ may interact with the basic surface of $FlhB_C$ and result in the export specificity switch (Figure 4.16A). There are two acidic loops in $FliK_C$ – loop 2 and loop 3. Loop 2 contains a highly conserved sequence $LxPx_1x_2LG$, where x_1 and x_2 are acidic residues (glutamic acid in *Salmonella*). Loop 3 also contains two acidic residues (aspartic acid in *Salmonella*), although these are not as well conserved (203).

These acidic residues were mutated to glycines to maintain loop structure while altering local chemistry (acidic to hydrophobic residues). The loss of these acidic residues showed no effect on complementation of swimming motility in a *fliK* null *Salmonella* strain. Cells expressing these *fliK* variants swam comparably to *fliK* null *Salmonella* cells expressing wild type *fliK* from recombinant plasmid (Figure 4.16B). This agrees with previously reported results that were published after these experiments were conducted (240). The combination of these results strong suggest the acidic loops are not involved in the export specificity switch, although they may still bind $FlhB_C$'s basic surface. This was not investigated in this study.

The goal of these experiments was to figure out how FliK may cause $FlhB_{CC}$ to dissociate from the rest of FlhB through direct interactions or enzymatic activity. What these experiments suggest is that FliK's role in the export specificity switch appears to be more complicated than a simple interaction. $FliK_C$ is stable, but presumably must be unfolded for export through the very narrow rod and hook channel. Does this unraveling expose binding sites covered in the structured C-terminal domain? That

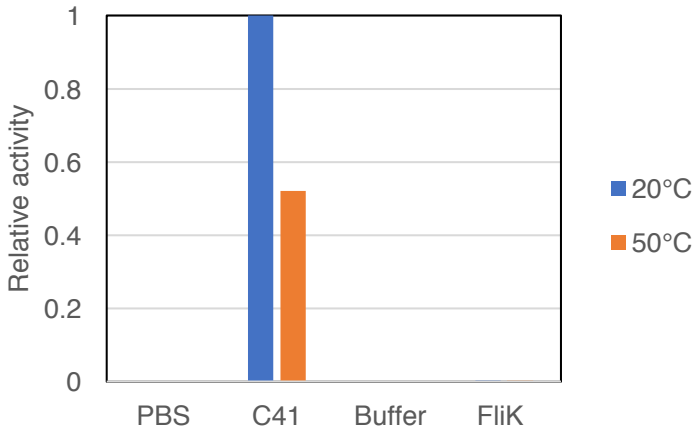


Figure 4.15 FliK is not a proline aminopeptidase

Proline aminopeptidase activity assay measured at A_{405} with 1 mM L-proline p-nitroanilide. PBS and Tris-saline buffer ("Buffer") served as absorbance blanks for assay. FliK in Tris-saline buffer is purified HexaHis-FliK in buffer at 6 μM. C41 soluble lysate in PBS ("C41") at 20°C is counted as 1.0 relative enzymatic activity.

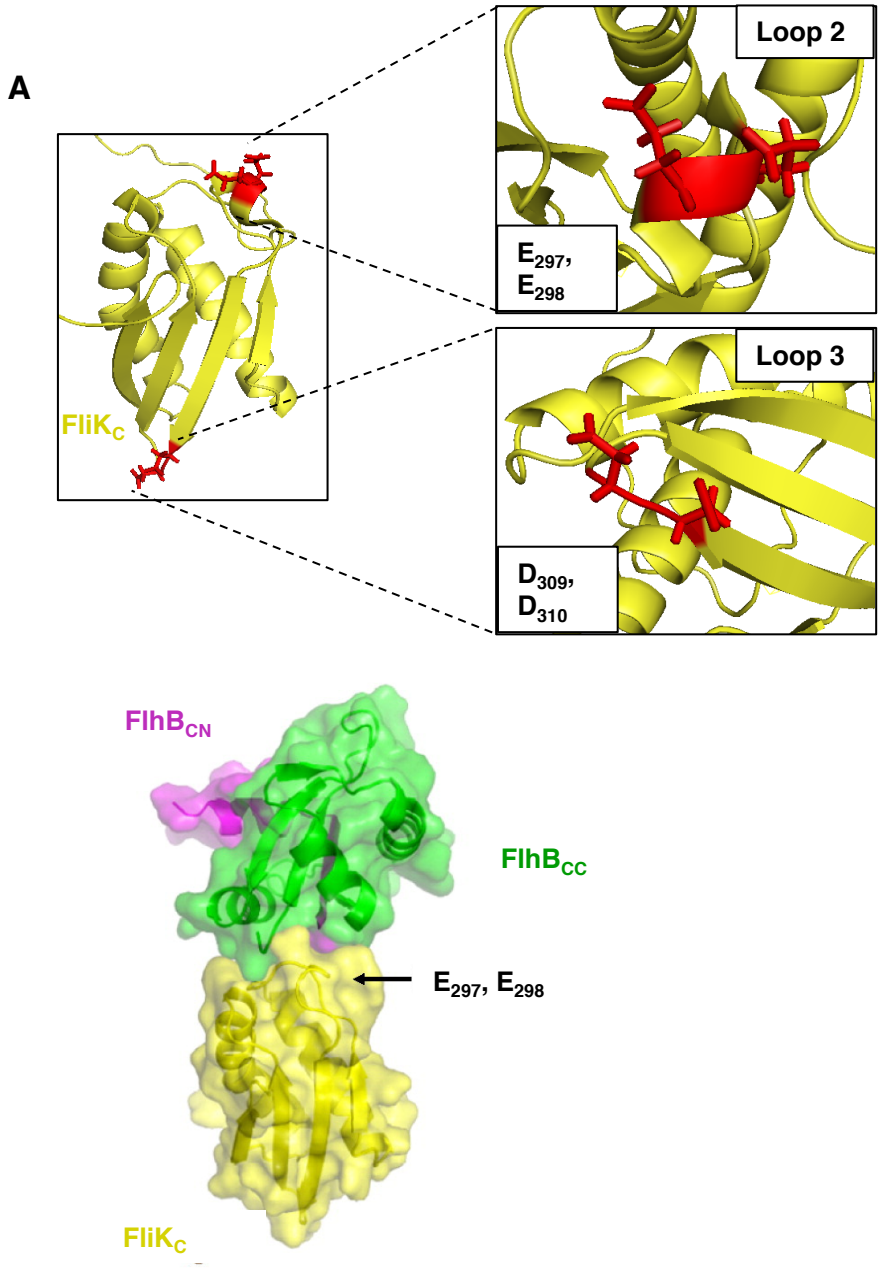


Figure 4.16 The acidic loops in $FliK_C$ are not responsible for causing the switch

(A) The core, structured domain of $FliK_C$ contains two α-helices, a β-sheet made up of four β-strands, and several loops. Some of these loops are acidic. $FliK_C$ loop 2 contains two glutamate (E) residues; $FliK_C$ loop 3 contains two aspartate (D) residues. A proposed model of $FliK_C$ acidic loop 2 binding to a basic cleft on $FlhB_C$ post-autocleavage is shown. Figure adapted from Mizuno et al., 2011.

(B) Complementation of SJW1103*ΔfliK::FRT* transformed with pBAD18amp plasmids carrying *fliK* wild type and acidic loop mutants on swimming motility agar (0.25% w/v agar) with 0.02% w/v L-arabinose after incubation for four hours at 37°C. Vector is empty pBAD18amp.

remains possible, and future work, proposed in the following discussion, should seek to address this question.

4.9 Discussion

$FlhB_C$ rapidly autocleaves into two cytosolic domains at the NPTH loop, with a half-life of approximately five minutes (190). These two domains, $FlhB_{CN}$ and $FlhB_{CC}$, can be co-purified in 1:1 stoichiometry, hinting that the non-covalent interactions holding the two domains together must be relatively strong. CD data investigating the melting temperature of autocleavable $FlhB_C$ and non-autocleavable $FlhB_C$ show that the protein is quite stable in either state, happy to exist as two tightly associated domains or one covalently-attached polypeptide with no discernable change in structure.

But, although structural work has provided evidence of possible non-covalent links between the two domains, an affinity between the $FlhB_{CN}$ and $FlhB_{CC}$ domains has not been established. $FlhB_{CC}$ was difficult to purify on its own, meaning that biophysical methods such as isothermal titration calorimetry (ITC) were not possible. $FlhB_{CN}$ could be purified without its $FlhB_{CC}$ partner domain. CD analysis showed that the protein did not adopt the expected, majority α-helical fold it shows when in association with $FlhB_{CC}$ in the published atomic resolution structure. The unusual structure of $FlhB_{CN}$ in isolation from $FlhB_{CC}$ makes studying the two domains separately difficult without the use of large tags, experiments which were not attempted in this study. *In vivo*, the deletion of $FlhB_{CC}$ from the chromosome can be partially rescued by overexpression from recombinant plasmid (190), suggesting that $FlhB_{CN}$ can adopt its native fold either while connected to the transmembrane domain of FlhB or upon association with the $FlhB_{CC}$ polypeptide in the cytosol. Because $FlhB_{CC}$ could not be purified without a large protein tag in this study, I did not attempt to mix $FlhB_{CN}$ and $FlhB_{CC}$ to see if the two domains could re-associate *in vitro*.

Since separating the two domains proved to be unfruitful, the autonomous switcher mutants were investigated as models to study the effect of the export specificity switch on the stability of $FlhB_C$. The two autonomous switcher mutations selected for this study, $S_{274}F$ and $R_{302}P$, are structurally critical based on analysis of the atomic resolution structure. S_{274} is located on β-strand$_2$ downstream of the primary

autocleavage site. R_{302} is centrally located in α-helix$_2$. These residues may also form a weak hydrogen bond in the core of the protein since they are at the limit that hydrogen bonding usually occurs. This may be one of the secondary, non-covalent bonds that hold FlhB together post-autocleavage.

CD analysis reveals that the autonomous switcher mutants are structurally similar to wild type FlhB, with one exception ($FlhB_C$ $R_{302}P$). The loss of the proposed hydrogen bond between S_{274} and R_{302} did not seem to have a major effect on FlhB's structure. It is difficult, then, to say what contribution this hydrogen bond makes on core arrangement. Presumably, it is one of many non-covalent bonds that keep the two polypeptides in association after autocleavage.

As previously discussed, changing R_{302} into a proline could damage α-helix$_2$ and lead to local unfolding, potentially exposing the core to the aqueous solution and therefore causing a major rearrangement of the folded domains to shield hydrophobic residues. This was the only variant of $FlhB_C$ that showed a shift in the CD data, indicating that some structural rearrangement had occurred. Unfortunately, it is not possible to analyse this shift further to determine what structural changes did occur. Crystallisation of $FlhB_C$ $R_{302}P$ was attempted, but crystals did not grow. Future work might purify $FlhB_C$ $R_{302}P$ and attempt other methods in solution or additional crystal screen conditions to characterise the structural changes caused by this point mutation.

Individually, the autonomous switcher mutants $S_{274}F$ and $R_{302}P$ have minor effects on the cell's ability to swim, both in the presence and absence of FliK. When combined as a double autonomous switcher mutant, however, these mutations are better able to rescue the swimming motility phenotype in the absence of FliK. These cells are inefficient swimmers, unable to outcompete wild type *Salmonella*, and only weakly express and export FliC.

This inefficient swimming is probably caused by a mistimed export specificity switch in the absence of FliK that results from sporadic autocleavage. The double autonomous switcher mutant undergoes inefficient autocleavage as observed *in vivo*, triggering the export specificity switch too early or too late. If the switch occurred too

early, a filament could begin to assemble too early. If the switch were delayed for too long, a filament may assemble on what would now be an elongated polyhook structure. In either case, the cell would not be able to swim as the short hook or long polyhook would not rotate and transfer torque efficiently.

Once the mutant FlhB does autocleave, the structural instability caused by the double autonomous switcher mutations results in the loss of $FlhB_{CC}$. What do these results tell us about the effect of the double autonomous switcher mutant on flagellar assembly? The double autonomous switcher mutations may serve to disrupt the interactions between $FlhB_{CN}$ and $FlhB_{CC}$ post-autocleavage just well enough to allow for the assembly of some competent flagella in a minority of cells, resulting in some rescue of the swimming phenotype in the absence of FliK.

Motility assays with the FlhB double autonomous switcher mutant in the presence or absence of FliK were conducted at 37°C to compare to the other FlhB variants. But the double autonomous switcher mutant was shown to be unstable during dialysis into low-salt buffer at 4°C (Figure 4.7). Future investigation could perform motility assays at higher temperatures, such as 42°C, to determine if this improves the cell's ability to swim. I hypothesise that the higher incubation temperature would lead to an earlier loss of the mutated $FlhB_{CC}$ from FlhB by promoting more rapid autocleavage of the double autonomous switcher mutant and then to a more efficient export specificity switch.

Williams and colleagues reported that cells with the autonomous switcher mutations swam slower than wild type, but ultimately produced motility rings on semi-solid agar similar to wild type. They measured these rates and ring sizes based only on the effective loss of FliK. In their strains, FliK was truncated to residues 1-300, out of the 405 residues in wild type FliK (243). In this study, the swimming rates of cells with autonomous switcher mutations were not measured, but the ring size is not near 50% of wild type *Salmonella* cells after 20 hours, which would have taken up the plate rather than swim just 30 mm as cells with the double autonomous switcher mutant did. The difference here could be the effective loss of FliK via a C-terminal truncation in the

Williams strains versus the near complete replacement of FliK with a kanamycin^R cassette in the strains used in this study.

So, what effects may the autonomous switcher mutants have on $FlhB_C$? The *in vitro* data show that the mutations greatly compromise the protein's thermal stability, with the exception of $W_{353}T$. $W_{353}T$ mimics a nonsense mutation by eliminating the last 30 residues of $FlhB_C$. Interestingly, these residues are not present in the atomic resolution structure, suggesting they are not well structured. Many of them are, however, assigned in the NMR spectra recorded for $FlhB_C$ and could adopt secondary structure in solution. The minor loss of thermal stability in the $W_{353}T$ mutant could be a result of the loss of this solution structure, lowering the energy barrier it takes to unfold the protein. These last 30 residues may also provide additional contacts between $FlhB_{CN}$ and $FlhB_{CC}$ by acting as a hinge between the two domains. The importance of these residues should be studied further to elucidate what role they may play in the protein's properties.

$S_{274}F$ has a greater effect than $W_{353}T$ on decreasing the melting temperature of $FlhB_C$. Replacing serine with phenylalanine removes a smaller, polar residue for a larger, aromatic one, but the phenylalanine would most likely stay confined to the protein's core. It may push out on the surrounding residues, particularly R_{302} on α-helix$_2$, and force them to adopt a slightly different conformation. This rearrangement and loss of a hydrogen bond may be what causes the decreased stability.

$R_{302}P$ results in the greatest decrease of the FlhB melting temperature that we could measure. Not only does a hydrophobic residue replace a surface exposed, charged residue, but prolines are known to break α-helices. The full effect of this mutation on α-helix$_2$ structure is unknown and should be studied further, but it is reasonable to suggest that the addition of a proline at least serves to destabilise the local region, remove a hydrogen bond partner, and increase the hydrophobicity on $FlhB_C$'s surface. This mutation almost certainly results in a greater structural change than $S_{274}F$, which would explain its greater effect on thermal stability.

The double autonomous switcher mutant, $S_{274}F$ $R_{302}P$, was not stable in the low-salt buffer needed for CD, either at 4°C or 25°C. The CD buffer (20 mM sodium phosphate buffer, pH 7.4, 100mM NaCl) was selected to minimise the background absorbance of buffer components during the CD experiments while maintaining $FlhB_C$ structure and approximating physiological conditions as much as possible. All other variants of $FlhB_C$ were stable in this buffer, so it is doubtful that the buffer caused the double autonomous switcher mutant to precipitate out of solution during dialysis. The double mutant was stable in the high salt (1 M NaCl), tris buffer used to purify proteins fused to the inteinCBD tag. The extra ionic charges in the high salt buffer may serve to stabilise the mutated $FlhB_C$'s structure.

What may happen with the double autonomous switcher mutant during dialysis? Unlike when the two domains are separately purified, and $FlhB_{CN}$ was shown to be super stable in solution, the two domains were expressed and purified *in vitro* simultaneously for the CD experiments. This should allow $FlhB_{CN}$ to adopt its native fold with the now unstable $FlhB_{CC}$ prior to or during autocleavage. As the double autonomous switcher mutant is stable in high salt, these two domains must stably interact long enough to be recovered after overnight cleavage from the expression inteinCBD tag. Once the salt concentration of the buffer was lowered though, $FlhB_{CC}$ would begin to destablise and could be pulling $FlhB_{CN}$ out of solution with it. This may explain why both domains can be seen in the precipitate.

In vitro stability is one matter. *In vivo*, where the protein exists in the crowded cell environment, is another. The addition of a 3xFLAG tag to the C-terminus of FlhB allows for the tracking of $FlhB_{CC}$'s fate in the cell. Under wild type conditions without FliK, and therefore the absence of an export specificity switch, FlhB autocleaves completely by late log phase, leaving a strong $FlhB_{CC}$ band. The double autonomous switcher mutant does not have a corresponding $FlhB_{CC}$ band in late log phase, but the protein has not completely autocleaved yet, as evidenced by the larger molecular weight band corresponding to the non-autocleavable variant of FlhB. This corroborates prior findings that these mutations lead to inefficient cleavage via an unknown mechanism (190).

The double autonomous switcher mutant has completely autocleaved once the cells have reached stationary phase. No $FlhB_{CC}$ band for the double autonomous switcher mutant was detected at this stage either, even though such a band could still be detected for wild type FlhB. This *in vivo* finding complements what we see *in vitro*. The double autonomous switcher mutant expressed well in the cell, but, once autocleavage occurred in an environment with low salt, like what might be found in a *Salmonella* cell, $FlhB_{CC}$ is not stable enough to maintain tertiary and secondary structure and is degraded in the cell.

The autonomous switcher mutants, especially the double autonomous switcher mutant, reveal the first clues about the downstream effects of FlhB autocleavage in the export specificity switch. These mutations destabilise $FlhB_C$ and ultimately cause the loss of $FlhB_{CC}$ from the rest of FlhB. Why might that be important? Early subunits bind to a surface-exposed hydrophobic pocket on $FlhB_{CC}$. Without this domain, early subunits would not have an export gate at the fT3SS. FlhB must autocleave, then, to allow for the proper timing of $FlhB_{CC}$ dissociation that leads to competent flagella.

This same effect is seen with wild type FlhB, but only in the presence of FliK. In the presence of FliK during log phase, $FlhB_{CC}$ starts to disappear. This change is not great, only a 2.4-fold difference, but this is probably due to the nature of flagella assembly in *Salmonella*. *Salmonella* is peritrichous, with many flagella at different stages of assembly per cell. In the wild type environment, it is possible this staggered assembly means that $FlhB_C$ has not yet switched at all of the fT3SS machines. The immunoblots analysed a population of cells, further complicating the picture by capturing both phases of flagellar assembly across a large population. Regardless, these results suggest that wild type $FlhB_{CC}$ dissociates from FlhB at the membrane in a FliK-dependent manner.

If two copies of FlhB are present at the fT3SS, it may mean one FlhB has “switched” while the other has not. This could explain why FliK is secreted intermittently through hook assembly. Hook length must be tightly controlled, but the range can still fall within 10% of average length (55 ± 6 nm). Two copies of FlhB could account for this range if both $FlhB_{CC}$’s would need to dissociate to trigger the switch.

Although $YscU_{CC}$ is shown to be secreted (198), we could not find similar evidence using 100 ml cultures of *Salmonella* expressing $FlhB_{Ct3xFLAG}$. We hypothesise that the export specificity switch activated by FliK simply leads to the loss of $FlhB_{CC}$ from the membrane. After dissociation, $FlhB_{CC}$ is either too unstable to be detected or is degraded by cellular proteases. Larger scale cultures may need to be studied to conclude if $FlhB_{CC}$ is secreted or not as it may also be unstable in the environment and, therefore, not detected by the anti-FLAG M2 antibody.

How $FliK_C$ may interact with $FlhB_C$ to trigger the dissociation of $FlhB_{CC}$ remains an open question. Genetic evidence is plentiful in both the flagellum and injectisome T3SSs and supports the hypothesis that $FliK_C/SctP_C$ is the export specificity switch domain (202). Much of the current biochemical and biophysical evidence for FliK's role in the switch uses full length protein and does not isolate this domain. This is problematic as FliK and SctP share N-terminal recognition motifs which bind FlhB and its homologues (86, 88, 89).

A few hypotheses of how FliK might cause the loss of $FlhB_{CC}$ were tested in this study. In 2011, Mizuno and colleagues proposed that two acidic loops found in the folded domain of $FliK_C$ may interact with the very basic surface of $FlhB_C$ (203). The result of the motility assay, where acidic residues in loops 2 and 3 were mutated to glycines, does not support the Mizuno hypothesis (Figure 4.16). This was confirmed by the corresponding author of the Mizuno paper at a recent T3SS conference and in a follow up publication measuring motility and hook length (240). It is now proposed that the conserved hydrophobic residues of loop 2 are critical for maintaining $FliK_C$ structure and, ultimately, for the export specificity switch. That hypothesis was not followed up in this study.

Biochemical and biophysical studies on full length FliK make it impossible to differentiate N-terminal and C-terminal interactions. Evans and colleagues showed that FliK could still bind to $FlhB_C$ *in vitro*, albeit very weakly, when its gate recognition motif is deleted (86). We attempted to use a $FliK_C$ truncation that includes the residues identified by genetic analysis as critical for the switch but does not include the N-terminal ruler and disordered linker domains of FliK. This truncation, $FliK_{270}$, was not

motile but was able to bind $FlhB_C$ *in vitro*. The lack of motility makes this truncation less desirable than a truncation which is export deficient but still activates the switch (204). For this reason, I focused on other hypotheses. But this truncation provides a valuable tool for follow-up investigations to study any direction interactions between the FliK T3S4 and $FlhB_C$.

Cell-based assays do suggest that $FlhB_{CC}$ dissociates from FlhB in a FliK-dependent manner. What about *in vitro*? To study this, C41 (DE3) soluble lysate from large scale expressions with empty recombinant plasmid or recombinant plasmid with full length FliK was incubated with $FlhB_{Ct3xFLAG}$. It was decided to use FliK soluble lysate with protease inhibitors to avoid purifying full length FliK, which often purifies with many breakdown products. Maintaining full length FliK in solution would more accurately simulate the cellular environment as well.

No loss of $FlhB_{CC}$ could be detected in this assay, but that may be due to two aspects of the experiment. The first, FliK was not as well expressed in the soluble fraction as expected. $FlhB_{Ct3xFLAG}$ (1 μM) was used to allow for detection, but this is much greater than the expected concentration of FlhB inside a cell. Future experiments should strive to increase the concentration of FliK while reducing the amount of FlhB used. Purifying FliK and incubating it, and its breakdown products, with FlhB may be another way to increase the amount of FliK used.

The second flaw of the experiment is a lack of pulling force. FliK, as an exported subunit, would cross the cell envelope and fold in the external environment. Assuming a direct interaction between $FliK_C$ and $FlhB_{CC}$, the force generated by FliK folding in the environment may serve to pull $FlhB_{CC}$ off from the rest of FlhB. Presumably, based on what we know of FlhB and $FlhB_{CC}$ interactions, this force would need to be strong enough to unravel at least a β-sheet. Furthermore, if FliK diffuses out of the cell rather than join a rate-constant subunit chain, the length of the hook would add a temporal component to the switch, as proposed by Erhardt and colleagues (142). For an immature, short hook, $FliK_N$ would start folding before $FliK_C$ could interact with $FlhB_{CC}$ and dissociate it from the rest of FlhB. Once the hook has reached a mature length, the diffusion of $FliK_N$ would be long enough to give $FliK_C$ the time to interact with

$FlhB_{CC}$. Once folding occurs, and the pulling force on $FlhB_{CC}$ grew strong enough, $FlhB_{CC}$ would then be pulled from the rest of FlhB. This model requires that $FliK_C$ and $FlhB_{CC}$ interact, which has yet to be settled convincingly.

Future studies must be conducted to conclude how $FliK_C$ interacts with $FlhB_{CC}$. That full length FliK interacts with $FlhB_{CC}$ is not in question, but many have assumed that $FliK_C$ must interact with $FlhB_{CC}$ based on the weight of genetic evidence that $FliK_C$ is critical to the switch. The simplest answer is that $FliK_C$ does interact with $FlhB_{CC}$, as suggested by the pulldown presented in this study using $FliK_{270}$, but that this interaction is not strong enough on its own to lead to the export specificity switch. A combination of biophysical techniques, evaluating the magnitude of the possible interaction between $FliK_C$ and $FlhB_{CC}$, and biochemical techniques such as scanning deletions in the $FliK_{linker}$ and $FliK_C$ domains could confirm important residues that may explain how FliK causes the dissociation of $FlhB_{CC}$ from the rest of FlhB. These variants of FliK should include a deletion of the FliK GRM to eliminate any possibility of background binding via the N-terminal secretion signal.

Until a direct interaction between the FliK T3S4 domain and $FlhB_C$ is conclusively shown, the prospect of another protein partner continues to be alluring. Based on kinetic data of FliK interactions with wild type and autocleavable-deficient FlhB and the results presented here, other proteins may play in a greater role in the export specificity switch than previously believed. Evidence points to FlhA or Flk in *Salmonella*, but there could also be unknown interactions that transmit the switching signal from FliK to $FlhB_C$ (191). These additional protein partners should be studied alongside FlhB and FliK while monitoring $FlhB_{CC}$ dissociation.

Chapter 5
Concluding remarks

5.1 Early subunit interactions with $FlhB_C$

Early subunits bind transiently to their export gate, FlhB (Figure 3.7). This lowers the forces and energy needed to remove subunit from the gate, allowing for rapid transit to the distal site of assembly. Export order of early subunits does not appear to be imposed entirely by affinity of their gate recognition motif (GRM) for the export gate, but, when subunits are separated into the two early substructures, there is a trend between subunit affinity for the export gate and export order. The earliest rod subunits, FliE and FlgB, bind with the greatest affinity while later rod subunits have weaker affinities. Furthermore, the hook subunit, FlgE, binds with a greater affinity than either the hook cap, FlgD, or the molecular ruler, FliK. Affinity of the GRM for the export gate may then be a result of both rough order of required assembly, as with FliE and FlgB needing to be exported and assembled first, or the number of subunits required for proper assembly, as would be the case with FlgE needing to outcompete FlgD and FliK to assemble a mature hook.

$FlhB_C$ contains a surface-exposed hydrophobic binding pocket, which is highly conserved across both species and FlhB/SctU homologues (Figure 3.19). This alignment, and the data presented here and in past publications, reinforces the importance of the interaction between early subunit GRM and the binding pocket on FlhB/SctU (86, 88, 89). Two residues in the binding pocket are of greatest importance based on chemical shift perturbations observed by NMR – $Alanine_{314}$ and $Leucine_{344}$. The other two residues identified in this binding pocket, $Alanine_{286}$ and $Proline_{287}$, may serve to contribute to the general hydrophobic environment but do not appear to be important in binding in either this or past studies. The caveat here being, of course, that prolines do not appear in 2D NMR probing for ^{1}H and ^{15}N, so P_{287} may have an unknown role in subunit binding. This is less likely, though, as the residues nearest $Alanine_{286}$ and $Proline_{287}$ do not experience chemical shift perturbations like the residues nearest $Alanine_{314}$ and $Leucine_{344}$ do, strongly suggesting that the early subunit GRM contacts $Alanine_{314}$ and $Leucine_{344}$.

NMR allowed for greater coverage of the FlhB amino acid sequence than crystallography did, especially in the C-terminal tail. Residues in this tail, which is missing from the atomic resolution structure, do experience chemical shift perturbations upon subunit binding. What role these residues play in subunit binding remains unknown. A solution structure could be solved and compared to the available atomic resolution structure to investigate differences in how FlhB crystallises and how it might fold in a cell-like environment.

Two of the autonomous switcher mutations are C-terminal tail truncations that match the cutoff of the atomic resolution structure folded domains ($Lysine_{351}$ and $Tryptophan_{353}$). Could this C-terminal tail have some role in stablising the interaction between $FlhB_{CC}$ and FlhB? NMR also allows for molecular dynamic measurements, data not collected in this study. These data could provide greater clarity in how $FlhB_{CC}$ and FlhB interact post-autocleavage by measuring how dynamic certain residues are. GRM peptide could also be added to monitor the change in dynamics after subunit binding and autonomous switcher mutations could be engineered in to observe how they affect individual residues. This information, in combination with a solution structure of $FlhB_C$, should help to explain the role of the C-terminal tail in early subunit binding and the export specificity switch.

How early subunit GRM sits in the binding pocket, and what structure it adapts while bound to the FlhB export gate, remain open questions. Is the GRM a helix while bound to $FlhB_C$ or is it completely unfolded by the time it associates with $FlhB_C$? How might the middle three residues, which vary greatly across subunits, contribute to GRM structure and binding? These questions can be investigated with techniques already known to work well with FlhB and early subunits such as *in vivo* cell- or population-based assays, ITC, and NMR. This data would also help to design novel therapeutic antagonists to subvert this interaction, leading to impaired flagellum or injectisome assembly. As both the GRM and subunit binding pocket are conserved across species, it is possible that therapeutics targeting both type III secretion systems in a range of pathogens could be found. Limiting cellular motility or the injection of effector molecules into the host cell could be a major step in treating some of the world's greatest bacterial threats.

5.2 A model for the flagellar export specificity switch

The results from this study provide evidence for the conclusion of the export specificity switch, namely the loss of $FlhB_{CC}$ from the rest of FlhB, demonstrating why autocleavage of $FlhB_C$ is vital to the switch. Data from this study and from others strongly support that the early subunit binding pocket is located on $FlhB_{CC}$. Once $FlhB_{CC}$ is removed from the FlhB export gate, the early subunits no longer have a gate to bind in preparation of export. How does the model for the switch look now? FliK is exported after interacting with $FlhB_{CC}$ via the FliK GRM. The unfolded N-terminus of FliK can then measure the nascent hook. If the hook is too short, the folding of the FliK N-terminus in the external environment pulls the rest of FliK from the export machinery without triggering the switch. But, if the hook has reached a mature length of 55 nm ± 6 nm, FliK will have the necessary time to trigger the loss of $FlhB_{CC}$ and lead to the export specificity switch (Figure 5.1).

This switch, presumably irreversible, means that the export machinery located at the base of that specific flagellum should now predominately export late subunits, starting with FlgM and continuing to filament export and assembly. Whether early subunits are still exported, especially in peritrichous organisms such as *Salmonella*, is difficult to discern due to staggered assembly of various flagella across a cell. But it is reasonable to assume that the export machinery moves permanently into late subunit export mediated by chaperone-subunit complexes and the FlhA export gate.

Data presented here suggest that the loss of a FlhB domain is FliK dependent, but how $FlhB_{CC}$ is removed by FliK from the rest of FlhB remains unclear. There is a wealth of genetic evidence on the importance of $FliK_C$ and its injectisome homologues to the export specificity switch, but biochemical data remains unclear and sparse. Further biochemical and biophysical analysis is needed to conclude how $FliK_C$ may activate the switch at $FlhB_C$, whether by direct interaction or the mediation of another protein. What about systems that lack FliK/SctP? What serves to trigger the loss of $FlhB_{CC}$/$SctU_{CC}$ in these T3SSs? These systems may have evolved different mechanisms to cause the loss of $FlhB_{CC}$/$SctU_{CC}$, such as less stable $FlhB_{CC}$ and $FlhB_{CN}$ interactions. They could also help uncover other T3SS subunits that may be involved in triggering the loss of $FlhB_{CC}$/$SctU_{CC}$.

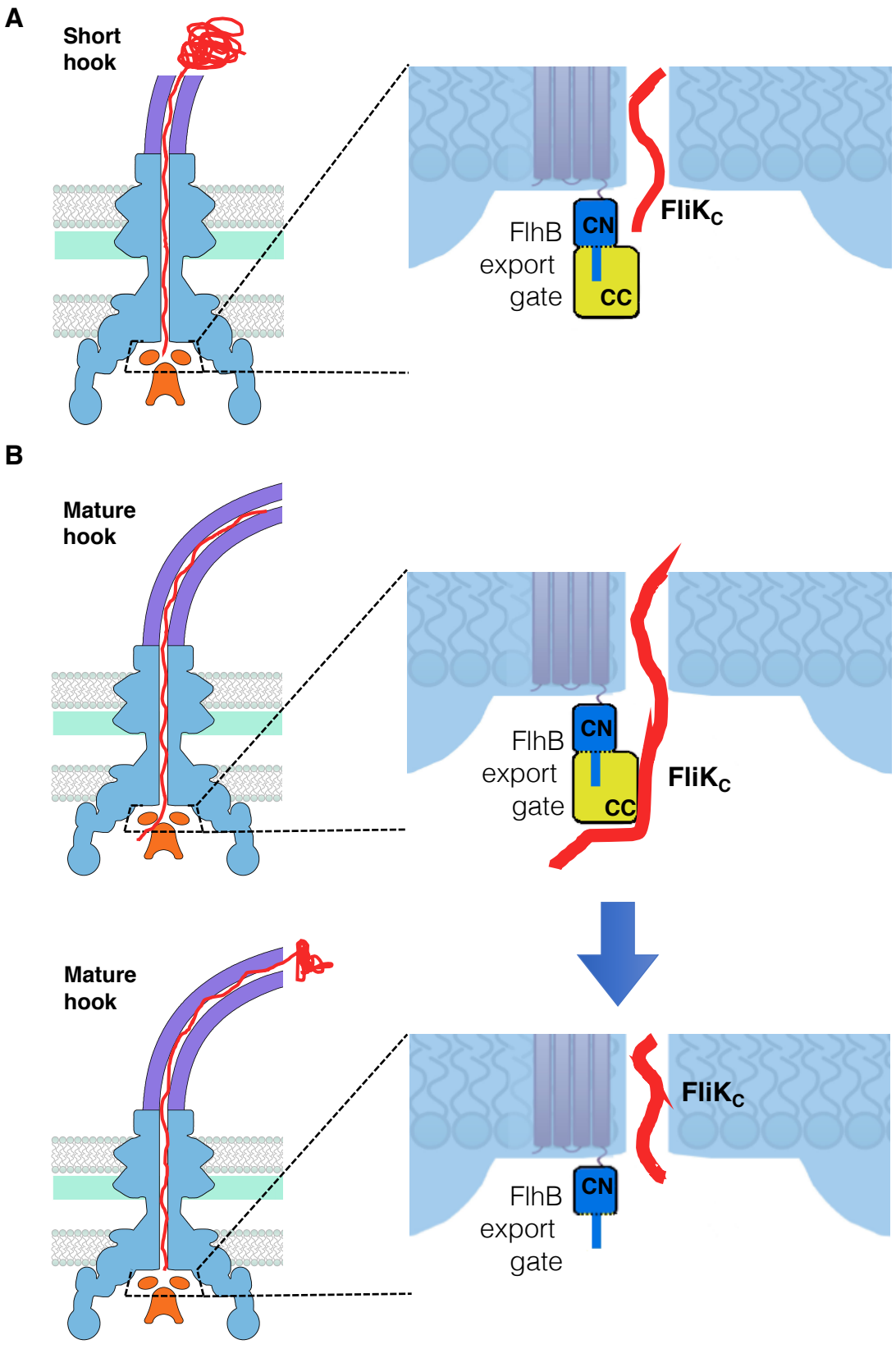


Figure 5.1 A proposed model for the export specificity switch at the fT3SS after the autocleavage of FlhB

FliK is secreted through the rod and hook after first binding to $FlhB_{CC}$ via the GRM.

(A) If $FliK_N$ measures a short hook, the N-terminus begins to fold in the environment and pulls $FliK_C$ out of the cell before causing the switch.

(B) If $FliK_N$ measures a mature hook, $FliK_C$ has enough time to cause $FlhB_{CC}$ to dissociate from the rest of FlhB via an unknown mechanism, triggering the switch. The loss of $FlhB_{CC}$ eliminates the export binding site for rod and hook subunits. The fT3SS will begin to secrete FlgM, leading to the expression of flagellin, and begin late subunit export and assembly.

Ongoing work in the Fraser lab also seeks to investigate whether FliK joins the subunit chain or is a chain breaker and exports via diffusion, which should help settle the question of temporal regulation of the export specificity switch. If FliK is a member of the constantly moving subunit chain, the pulling force generated by FliK as it folds in the environment would need to be enough to pull $FlhB_{CC}$ from FlhB, either directly or via another unknown protein.

Secretion of $SctU_{CC}$ in the *Yersinia* injectisome is thought to be important to activate the export specificity switch in the injectisome (198). I could find no evidence that $FlhB_{CC}$ is secreted, although the loss of a subunit binding site is consistent. I hypothesise that $FlhB_{CC}$ is too unstable to maintain structure without $FlhB_{CN}$ and is quickly degraded by cellular proteases after dissociation. This observation highlights a key difference between the flagellar and injectisome T3SS. Furthermore, the loss of *Yersinia* $SctU_{CC}$ to activate the export specificity switch was not shown to be SctP-dependent, although the evidence suggests that this is the case (198).

The flagellar and injectisome export machinery may diverge slightly on the export specificity switch mechanism, the loss of the early subunit binding site on $FlhB_{CC}$/$SctU_{CC}$ is a critical junction between export phases shared by both systems. The loss of $FlhB_{CC}$/$SctU_{CC}$ should continue to be studied in other organisms, but current data strongly supports that this mechanism is one shared across type III secretion systems.

Whether used to move through an environment or to associate with hosts as pathogens or symbionts, type III secretion systems are a powerful tool utilised by many bacteria. Ordered export by type III secretion machines is one of the most elegant examples of multi-step precision control that combines gene expression, subunit export from the cell, and protein subunit assembly into the nascent nanomachine. FlhB plays a critical role in imposing order onto subunit export. It serves not only as the export gate for early subunits but also to ensure the fT3SS switches export specificity at the appropriate time by sacrificing part of itself. Although the exact mechanism of the export specificity switch remains unclear, the work presented here contributes new pieces to our understanding of how biological nanomachines operate.

References

1. Pallen MJ, Gophna U. 2007. Bacterial Flagella and Type III Secretion: Case Studies in the Evolution of Complexity, p. 30–47. Gene and Protein Evolution. KARGER, Basel.
2. Terashima H, Kojima S, Homma M. 2008. Chapter 2 Flagellar Motility in Bacteria. Structure and Function of Flagellar MotorInternational Review of Cell and Molecular Biology. Elsevier Inc.
3. Berg HC, Anderson RA. 1973. Bacteria Swim by Rotating their Flagellar Filaments. Nature 245:380–382.
4. Armitage JP, Macnab RM. 1987. Unidirectional, intermittent rotation of the flagellum of *Rhodobacter sphaeroides*. J Bacteriol 169:514–518.
5. Schmitt R. 2002. Sinorhizobial chemotaxis: A departure from the enterobacterial paradigm. Microbiology 148:627–631.
6. Szurmant H, Ordal GW. 2004. Diversity in Chemotaxis Mechanisms among the Bacteria and Archaea. Microbiol Mol Biol Rev 68:301–319.
7. Armitage JP. 1999. Bacterial tactic responses. Adv Microb Physiol 41:229–89.
8. Welch M, Oosawa K, Aizawa SI, Eisenbach M. 1993. Phosphorylation-dependent binding of a signal molecule to the flagellar switch of bacteria. Proc Natl Acad Sci 90:8787–8791.
9. Asai Y, Kojima S, Kato H, Nishioka N, Kawagishi I, Homma M. 1997. Putative channel components for the fast-rotating sodium-driven flagellar motor of a marine bacterium. J Bacteriol 179:5104–5110.
10. Dean GE, Macnab RM, Stader J, Matsumura P, Burks C. 1984. Gene sequence and predicted amino acid sequence of the *motA* protein, a membrane-associated protein required for flagellar rotation in *Escherichia coli*. J Bacteriol 159:991–999.
11. Stader J, Matsumura P, Vacante D, Dean GE, Macnab RM. 1986. Nucleotide sequence of the *Escherichia coli motB* gene and site-limited incorporation of its product into the cytoplasmic membrane. J Bacteriol 166:244–252.
12. Kojima M, Kubo R, Yakushi T, Homma M, Kawagishi I. 2007. The bidirectional polar and unidirectional lateral flagellar motors of *Vibrio alginolyticus* are controlled by a single CheY species. Mol Microbiol 64:57–67.
13. Asai Y, Yakushi T, Kawagishi I, Homma M. 2003. Ion-coupling Determinants of

Na+-driven and H+-driven Flagellar Motors. J Mol Biol 327:453–463.

14. Kojima S, Blair DF. 2004. Solubilization and purification of the MotA/MotB complex of *Escherichia coli*. Biochemistry 43:26–34.
15. Reid SW, Leake MC, Chandler JH, Lo C-J, Armitage JP, Berry RM. 2006. The maximum number of torque-generating units in the flagellar motor of *Escherichia coli* is at least 11. Proc Natl Acad Sci U S A 103:8066–8071.
16. Leake MC, Chandler JH, Wadhams GH, Bai F, Berry RM, Armitage JP. 2006. Stoichiometry and turnover in single, functioning membrane protein complexes. Nature 443:355–358.
17. Lloyd SA, Blair DF. 1997. Charged residues of the rotor protein FliG essential for torque generation in the flagellar motor of *Escherichia coli*. J Mol Biol 266:733–44.
18. Zhou J, Blair DF. 1997. Residues of the cytoplasmic domain of MotA essential for torque generation in the bacterial flagellar motor. J Mol Biol 273:428–439.
19. Zhou J, Lloyd SA, Blair DF. 1998. Electrostatic interactions between rotor and stator in the bacterial flagellar motor. Proc Natl Acad Sci U S A 95:6436–6441.
20. Blair DF, Berg HC. 1990. The MotA protein of *E. coli* is a proton-conducting component of the flagellar motor. Cell 60:439–449.
21. Stolz B, Berg HC. 1991. Evidence for interactions between MotA and MotB, torque-generating elements of the flagellar motor of *Escherichia coli*. J Bacteriol 173:7033–7037.
22. Zhou J, Sharp LL, Tang HL, Lloyd SA, Billings S, Braun TF, Blair DF. 1998. Function of Protonatable Residues in the Flagellar Motor of *Escherichia coli*: a Critical Role for Asp 32 of MotB. J Bacteriol 180:2729–2735.
23. Che YS, Nakamura S, Kojima S, Kami-ike N, Namba K, Minamino T. 2008. Suppressor analysis of the MotB(D33E) mutation to probe bacterial flagellar motor dynamics coupled with proton translocation. J Bacteriol 190:6660–6667.
24. Mot R, Vanderleyden J. 1994. The C-terminal sequence conservation between OmpA-related outer membrane proteins and MotB suggests a common function in both Gram-positive and Gram-negative bacteria, possibly in the interaction of these domains with peptidoglycan. Mol Microbiol 12:333–334.
25. Koebnik R. 1995. Proposal for a peptidoglycan-associating alpha-helical motif in the C-terminal regions of some bacterial cell-surface proteins. Mol Microbiol

16:1269–1270.

26. Ryu WS, Berry RM, Berg HC. 2000. Torque-generating units of the flagellar motor of *Escherichia coli* have a high duty ratio. Nature 403:444–447.
27. Muramoto K, Kawagishi I, Kudo S, Magariyama Y, Imae Y, Homma M. 1995. High-speed rotation and speed stability of the sodium-driven flagellar motor in *Vibrio alginolyticus*. J Mol Biol 251:50–58.
28. Magariyama Y, Sugiyama S, Kudo S. 2001. Bacterial swimming speed and rotation rate of bundled flagella. FEMS Microbiol Lett 199:125–129.
29. Magariyama Y, Sugiyama S, Muramoto K, Kawagishi I, Imae Y, Kudo S. 1995. Simultaneous measurement of bacterial flagellar rotation rate and swimming speed. Biophys J 69:2154–62.
30. Atsumi T, Maekawa Y, Yamada T, Kawagishi I, Imae Y, Homma M. 1996. Effect of viscosity on swimming by the lateral and polar flagella of *Vibrio alginolyticus*. J Bacteriol 178:5024–6.
31. Guttenplan SB, Shaw S, Kearns DB. 2013. The cell biology of peritrichous flagella in *Bacillus subtilis*. Mol Microbiol 87:211–229.
32. Dasgupta N, Arora SK, Ramphal R. 2000. *fleN*, a gene that regulates flagellar number in *Pseudomonas aeruginosa*. J Bacteriol 182:357–364.
33. Pandza S, Baetens M, Park CH, Au T, Keyhan M, Matin a. 2000. The G-protein FlhF has a role in polar flagellar placement and general stress response induction in *Pseudomonas putida*. Mol Microbiol 36:414–423.
34. Murray TS, Kazmierczak BI. 2006. FlhF Is required for swimming and swarming in *Pseudomonas aeruginosa*. J Bacteriol 188:6995–7004.
35. Balaban M, Hendrixson DR. 2011. Polar flagellar biosynthesis and a regulator of flagellar number influence spatial parameters of cell division in *Campylobacter jejuni*. PLoS Pathog 7:18–25.
36. Correa NE, Peng F, Klose KE. 2005. Roles of the regulatory proteins FlhF and FlhG in the *Vibrio cholerae* flagellar transcription hierarchy. J Bacteriol 187:6324–6332.
37. Chaban B, Hughes HV, Beeby M. 2015. The flagellum in bacterial pathogens: For motility and a whole lot more. Semin Cell Dev Biol 46:91–103.
38. Aihara E, Closson C, Matthis AL, Schumacher MA, Engevik AC, Zavros Y, Ottemann KM, Montrose MH. 2014. Motility and Chemotaxis Mediate the

Preferential Colonization of Gastric Injury Sites by *Helicobacter pylori*. PLoS Pathog 10.

39. Bolton DJ. 2015. *Campylobacter* virulence and survival factors. Food Microbiol 48:99–108.
40. Stecher B, Hapfelmeier S, Müller C, Kremer M, Stallmach T, Hardt W. 2004. Flagella and chemotaxis are required for efficient induction of *Salmonella enterica* serovar Typhimurium colitis in streptomycin-pretreated mice. Infect Immun 72:4138–50.
41. Rossez Y, Wolfson EB, Holmes A, Gally DL, Holden NJ. 2015. Bacterial flagella: twist and stick, or dodge across the kingdoms. PLoS Pathog 11:e1004483.
42. Girón JA, Torres AG, Freer E, Kaper JB. 2002. The flagella of enteropathogenic *Escherichia coli* mediate adherence to epithelial cells. Mol Microbiol 44:361–79.
43. Roy K, Hilliard GM, Hamilton DJ, Luo J, Ostmann MM, Fleckenstein JM. 2009. Enterotoxigenic *Escherichia coli* EtpA mediates adhesion between flagella and host cells. Nature 457:594–8.
44. McSweegan E, Walker RI. 1986. Identification and characterization of two *Campylobacter jejuni* adhesins for cellular and mucous substrates. Infect Immun 53:141–148.
45. Lillehoj EP, Kim BT, Kim KC. 2002. Identification of *Pseudomonas aeruginosa* flagellin as an adhesin for Muc1 mucin. Am J Physiol Lung Cell Mol Physiol 282:L751--6.
46. Wolfson E (badass). 2013. Molecular and functional characterisation of the adherent properties of H7 flagella. University of Edinburgh.
47. Cossart P, Sansonetti PJ. 2004. Bacterial invasion: the paradigms of enteroinvasive pathogens. Science 304:242–8.
48. Eriksson S, Lucchini S, Thompson A, Rhen M, Hinton JCD. 2003. Unravelling the biology of macrophage infection by gene expression profiling of intracellular *Salmonella enterica*. Mol Microbiol 47:103–118.
49. Mahajan A, Currie CG, Mackie S, Tree J, Mcateer S, Mckendrick I, Mcneilly TN, Roe A, La Ragione RM, Woodward MJ, Gally DL, Smith DGE. 2009. An investigation of the expression and adhesin function of H7 flagella in the

interaction of *Escherichia coli* O157: H7 with bovine intestinal epithelium. Cell Microbiol 11:121–137.

50. Ott M, Messner P, Heesemann J, Marre R, Hacker J. 1991. Temperature-dependent expression of flagella in *Legionella*. J Gen Microbiol 137:1955–1961.
51. Andrewes FW. 1922. Studies in group-agglutination I. The *Salmonella* group and its antigenic structure. J Pathol Bacteriol 25:505–521.
52. Bonifield HR, Hughes KT. 2003. Flagellar Phase Variation in *Salmonella enterica* Is Mediated by a Posttranscriptional Control Mechanism. J Bacteriol 185:3567–3574.
53. Stewart MK, Cookson BT. 2012. Non-genetic diversity shapes infectious capacity and host resistance. Trends Microbiol 20:461–466.
54. Andersen-Nissen E, Smith KD, Strobe KL, Barrett SL, Cookson BT, Logan SM, Aderem A. 2005. Evasion of Toll-like receptor 5 by flagellated bacteria. Proc Natl Acad Sci U S A 102:9247–9252.
55. World Health Organization. 2016. *Salmonella* (non-typhoidal) Fact Sheet.
56. Centers for Disease Control and Prevention. 2016. National Antimicrobial Resistance Monitoring System (NARMS): 2014 Human Isolates Surveillance Report.
57. Centers for Disease Control and Prevention (CDC). 2016. National Enteric Disease Surveillance: *Salmonella* Annual Report, 2013.
58. Yamaguchi S, Fujita H, Sugata K, Taira T, Iino T. 1984. Genetic analysis of H2, the structural gene for phase-2 flagellin in *Salmonella*. J Gen Microbiol 130:255–65.
59. Costa TRD, Felisberto-Rodrigues C, Meir A, Prevost MS, Redzej A, Trokter M, Waksman G. 2015. Secretion systems in Gram-negative bacteria: structural and mechanistic insights. Nat Rev Microbiol 13:343–359.
60. Abby SS, Cury J, Guglielmini J, Néron B, Touchon M, Rocha EPC. 2016. Identification of protein secretion systems in bacterial genomes. Sci Rep 6:23080.
61. Korotkov K V., Sandkvist M, Hol WGJ. 2012. The type II secretion system: biogenesis, molecular architecture and mechanism. Nat Rev Microbiol 10:336–351.

62. Nivaskumar M, Francetic O. 2014. Type II secretion system: A magic beanstalk or a protein escalator. Biochim Biophys Acta - Mol Cell Res 1843:1568–1577.
63. Trokter M, Felisberto-Rodrigues C, Christie PJ, Waksman G. 2014. Recent advances in the structural and molecular biology of type IV secretion systems. Curr Opin Struct Biol 27:16–23.
64. Eicher T, Seeger MA, Anselmi C, Zhou W, Brandstätter L, Verrey F, Diederichs K, Faraldo-Gómez JD, Pos KM. 2014. Coupling of remote alternating-access transport mechanisms for protons and substrates in the multidrug efflux pump AcrB. Elife 3:1–26.
65. Junker M, Besingi RN, Clark PL. 2009. Vectorial transport and folding of an autotransporter virulence protein during outer membrane secretion. Mol Microbiol 71:1323–1332.
66. Bavro VN, Pietras Z, Furnham N, Pérez-Cano L, Fernández-Recio J, Pei XY, Misra R, Luisi B. 2008. Assembly and Channel Opening in a Bacterial Drug Efflux Machine. Mol Cell 30:114–121.
67. Loquet A, Sgourakis NG, Gupta R, Giller K, Riedel D, Goosmann C, Griesinger C, Kolbe M, Baker D, Becker S, Lange A. 2012. Atomic model of the type III secretion system needle. Nature 486:276–279.
68. Radics J, Königsmaier L, Marlovits TC. 2014. Structure of a pathogenic type 3 secretion system in action. Nat Struct Mol Biol 21:82–87.
69. Hultgren SJ, Normark S, Abraham SN. 1991. Chaperone-assisted assembly and molecular architecture of adhesive pili. Annu Rev Microbiol 45:383–415.
70. Vetsch M, Puorger C, Spirig T, Grauschopf U, Weber-Ban EU, Glockshuber R. 2004. Pilus chaperones represent a new type of protein-folding catalyst. Nature 431:329–333.
71. Galan J, Wolf-Watz H. 2006. Protein delivery into eukaryotic cells by type III secretion machines. Nature 444:567–573.
72. Cornelis G. 2006. The type III secretion injectisome. Nat Rev Microbiol 4:811–825.
73. Preston GM. 2007. Metropolitan Microbes: Type III Secretion in Multihost Symbionts. Cell Host Microbe 2:291–294.
74. Diepold A, Armitage JP. 2015. Type III secretion systems: the bacterial

flagellum and the injectisome. Philos Trans R Soc B 370:20150020.

75. Akeda Y, Galán JE. 2005. Chaperone release and unfolding of substrates in type III secretion. Nature 437:911–5.
76. Evans LDB, Stafford GP, Ahmed S, Fraser GM, Hughes C. 2006. An escort mechanism for cycling of export chaperones during flagellum assembly. Proc Natl Acad Sci U S A 103:17474–9.
77. Kuwajima G, Kawagishi I, Homma M, Asaka J, Kondo E, Macnab RM. 1989. Export of an N-terminal fragment of *Escherichia coli* flagellin by a flagellum-specific pathway. Proc Natl Acad Sci U S A 86:4953–4957.
78. Sory MP, Boland A, Lambermont I, Cornelis GR. 1995. Identification of the YopE and YopH domains required for secretion and internalization into the cytosol of macrophages, using the *cyaA* gene fusion approach. Proc Natl Acad Sci U S A 92:11998–2002.
79. Chilcott GS, Hughes KT. 1998. The type III secretion determinants of the flagellar anti-transcription factor, FlgM, extend from the amino-terminus into the anti-sigma28 domain. Mol Microbiol 30:1029–1040.
80. Namba K. 2001. Roles of partly unfolded conformations in macromolecular self-assembly. Genes to Cells 6:1–12.
81. Lloyd SA, Norman M, Rosqvist R, Wolf-Watz H. 2001. *Yersinia* YopE is targeted for type III secretion by N-terminal, not mRNA, signals. Mol Microbiol 39:520–531.
82. Warren SM, Young GM. 2005. An Amino-Terminal Secretion Signal Is Required for YplA Export by the Ysa, Ysc, and Flagellar Type III Secretion Systems of *Yersinia enterocolitica* Biovar 1B. J Bacteriol 187:6075–6083.
83. Karavolos MH, Wilson M, Henderson J, Lee JJ, Khan CMA. 2005. Type III Secretion of the *Salmonella* Effector Protein SopE Is Mediated via an N-Terminal Amino Acid Signal and Not an mRNA Sequence. J Bacteriol 187:1559–1567.
84. Végh BM, Gál P, Dobó J, Závodszky P, Vonderviszt F. 2006. Localization of the flagellum-specific secretion signal in *Salmonella* flagellin. Biochem Biophys Res Commun 345:93–98.
85. Stafford GP, Evans LDB, Krumscheid R, Dhillon P, Fraser GM, Hughes C. 2007. Sorting of early and late flagellar subunits after docking at the

membrane ATPase of the type III export pathway. J Mol Biol 374:877–82.
86. Evans LDB, Poulter S, Terentjev EM, Hughes C, Fraser GM. 2013. A chain mechanism for flagellum growth. Nature 504:287–90.
87. Singer HM, Erhardt M, Hughes KT. 2014. Comparative analysis of the secretion capability of early and late flagellar type III secretion substrates. Mol Microbiol 93:505–520.
88. Bergeron JRC, Fernández L, Wasney GA, Vuckovic M, Reffuveille F, Hancock REW, Strynadka NCJ. 2016. The structure of a type 3 secretion system (T3SS) ruler protein suggests a molecular mechanism for needle length sensing. J Biol Chem 291:1676–1691.
89. Ho O, Rogne P, Edgren T, Wolf-Watz H, Login FH, Wolf-Watz M. 2017. Characterization of the Ruler Protein Interaction Interface on the Substrate Specificity Switch Protein in the *Yersinia* Type III Secretion System. J Biol Chem 292:3299–3311.
90. Minamino T, Namba K. 2008. Distinct roles of the FliI ATPase and proton motive force in bacterial flagellar protein export. Nature 451:485–8.
91. Paul K, Erhardt M, Hirano T, Blair DF, Hughes KT. 2008. Energy source of flagellar type III secretion. Nature 451:489–92.
92. Erhardt M, Namba K, Hughes KT. 2010. Bacterial nanomachines: the flagellum and type III injectisome. Cold Spring Harb Perspect Biol 2:1–22.
93. Abby SS, Rocha EPC. 2012. The Non-Flagellar Type III Secretion System Evolved from the Bacterial Flagellum and Diversified into Host-Cell Adapted Systems. PLoS Genet 8.
94. Resistance TR on A. 2016. Tackling Drug-Resistant Infections Globally: Final Report and Recommendations.
95. Anantharajah A, Mingeot-Leclercq MP, Van Bambeke F. 2016. Targeting the Type Three Secretion System in *Pseudomonas aeruginosa*. Trends Pharmacol Sci 37:734–749.
96. Kerridge D. 1960. The effect of inhibitors on the formation of flagella by *Salmonella typhimurium*. J Gen Microbiol 23:519–538.
97. Goh YS, Armour KL, Clark MR, Grant AJ, Mastroeni P. 2016. Igg Subclasses Targeting the Flagella of *Salmonella enterica* serovar Typhimurium Can Mediate Phagocytosis and Bacterial Killing. J Vaccines Vaccin 7:1–15.

98. Rasmussen L, White EL, Pathak A, Ayala JC, Wang H, Wu JH, Benitez JA, Silva AJ. 2011. A high-throughput screening assay for inhibitors of bacterial motility identifies a novel inhibitor of the Na+-driven flagellar motor and virulence gene expression in *Vibrio cholerae*. Antimicrob Agents Chemother 55:4134–4143.
99. Renault TT, Abraham AO, Bergmiller T, Paradis G, Rainville S, Charpentier E, Guet CC, Tu Y, Namba K, Keener JP, Minamino T, Erhardt M. 2017. Bacterial flagella grow through an injection-diffusion mechanism. Elife 6.
100. Robson A, Gold V a M, Hodson S, Clarke AR, Collinson I. 2009. Energy transduction in protein transport and the ATP hydrolytic cycle of SecA. Proc Natl Acad Sci U S A 106:5111–5116.
101. Lenders MHH, Beer T, Smits SHJ, Schmitt L. 2016. In vivo quantification of the secretion rates of the hemolysin A Type I secretion system. Sci Rep 6:33275.
102. Singer HM, Erhardt M, Steiner AM, Zhang MM, Yoshikami D, Bulaj G, Olivera BM, Hughes KT. 2012. Selective purification of recombinant neuroactive peptides using the flagellar type III secretion system. MBio 3:1–9.
103. Dobó J, Varga J, Sajó R, Végh BM, Péter, Gal PZ, Vonderviszt F. 2010. Application of a short, disordered n-terminal flagellin segment, a fully functional flagellar type III export signal, to expression of secreted proteins. Appl Environ Microbiol 76:891–899.
104. Frye J, Karlinsey JE, Felise HR, Marzolf B, Dowidar N, McClelland M, Hughes KT. 2006. Identification of New Flagellar Genes of *Salmonella enterica* serovar Typhimurium. J Bacteriol 188:2233–2243.
105. Chevance FF V, Hughes KT. 2008. Coordinating assembly of a bacterial macromolecular machine. Nat Rev Microbiol 6:455–65.
106. Wang S, Fleming RT, Westbrook EM, Matsumura P, McKay DB. 2006. Structure of the *Escherichia coli* FlhDC complex, a prokaryotic heteromeric regulator of transcription. J Mol Biol 355:798–808.
107. Komeda Y, Suzuki H, Ishidsu J ichi, Iino T. 1975. The role of cAMP in flagellation of *Salmonella typhimurium*. MGG Mol Gen Genet 142:289–298.
108. Kelly A, Goldberg MD, Carroll RK, Danino V, Hinton JCD, Dorman CJ. 2004. A global role for Fis in the transcriptional control of metabolism and type III secretion in *Salmonella enterica* serovar Typhimurium. Microbiology

150:2037–2053.

109. Singer HM, Erhardt M, Hughes KT. 2013. RflM functions as a transcriptional repressor in the autogenous control of the *Salmonella* flagellar master operon flhDC. J Bacteriol 195:4274–4282.
110. Mouslim C, Hughes KT. 2014. The Effect of Cell Growth Phase on the Regulatory Cross-Talk between Flagellar and SPI1 Virulence Gene Expression. PLoS Pathog 10.
111. Tomoyasu T, Takaya A, Isogai E, Yamamoto T. 2003. Turnover of FlhD and FlhC, master regulator proteins for *Salmonella* flagellum biogenesis, by the ATP-dependent ClpXP protease. Mol Microbiol 48:443–452.
112. Liu X, Matsumura P. 1994. The FlhD/FlhC complex, a transcriptional activator of the *Escherichia coli* flagellar class II operons. J Bacteriol 176:7345–51.
113. Liu X, Fujita N, Ishihama A, Matsumura P. 1995. The C-terminal region of the alpha subunit of *Escherichia coli* RNA polymerase is required for transcriptional activation of the flagellar level II operons by the FlhD/FlhC complex. J Bacteriol 177:5186–5188.
114. Ohnishi K, Kutsukake K, Suzuki H, Iino T. 1990. Gene *fliA* encodes an alternative sigma factor specific for flagellar operons in *Salmonella* typhimurium. MGG Mol Gen Genet 221:139–147.
115. Karlinsey JE, Tanaka S, Bettenworth V, Yamaguchi S, Boos W, Aizawa SI, Hughes KT. 2000. Completion of the hook-basal body complex of the *Salmonella typhimurium* flagellum is coupled to FlgM secretion and *fliC* transcription. Mol Microbiol 37:1220–1231.
116. Ohnishi K, Kutsukake K, Suzuki H, Lino T. 1992. A novel transcriptional regulation mechanism in the flagellar regulon of Salmonella typhimurium: an anti-sigma factor inhibits the activity of the flagellum-specific Sigma factor, σ^F. Mol Microbiol 6:3149–3157.
117. Chadsey MS, Karlinsey JE, Hughes KT. 1998. The flagellar anti-sigma factor FlgM actively dissociates *Salmonella typhimurium* sigma 28 RNA polymerase holoenzyme. Genes Dev 12:3123–3136.
118. Chadsey MS, Hughes KT. 2001. A multipartite interaction between *Salmonella* transcription factor σ^{28} and its anti-sigma factor FlgM: implications for σ^{28} holoenzyme destabilization through stepwise binding. J Mol Biol 306:915–929.

119. Bennett JCQ, Thomas J, Fraser GM, Hughes C. 2001. Substrate complexes and domain organization of the *Salmonella* flagellar export chaperones FlgN and FliT. Mol Microbiol 39:781–791.
120. Yamamoto S, Kutsukake K. 2006. FliT acts as an anti-$FlhD_2C_2$ factor in the transcriptional control of the flagellar regulon in *Salmonella enterica* serovar Typhimurium. J Bacteriol 188:6703–6708.
121. Hung CC, Haines L, Altier C. 2012. The flagellar regulator *fliT* represses *Salmonella* Pathogenicity Island 1 through *flhDC* and *fliZ*. PLoS One 7.
122. Lucas RL, Lostroh CP, DiRusso CC, Spector MP, Wanner BL, Lee CA. 2000. Multiple factors independently regulate *hilA* and invasion gene expression in *Salmonella enterica* serovar Typhimurium. J Bacteriol 182:1872–82.
123. Kutsukake K, Ohya Y, Iino T. 1990. Transcriptional analysis of the flagellar regulon of *Salmonella typhimurium*. J Bacteriol 172:741–747.
124. Chilcott GS, Hughes KT. 2000. Coupling of Flagellar Gene Expression to Flagellar Assembly in *Salmonella enterica* serovar Typhimurium and Escherichia coli. Microbiol Mol Biol Rev 64:694–708.
125. Macnab RM. 2003. How bacteria assemble flagella. Annu Rev Microbiol 57:77–100.
126. Evans LDB, Hughes C, Fraser GM. 2014. Building a flagellum outside the bacterial cell. Trends Microbiol 22:566–572.
127. Yonekura K, Maki-Yonekura S, Namba K. 2003. Complete atomic model of the bacterial flagellar filament by electron cryomicroscopy. Nature 424:643–50.
128. Fujii T, Kato T, Hiraoka KD, Miyata T, Minamino T, Chevance FF V., Hughes KT, Namba K. 2017. Identical folds used for distinct mechanical functions of the bacterial flagellar rod and hook. Nat Commun 8:14276.
129. Zhao X, Zhang K, Boquoi T, Hu B, Motaleb MA, Miller KA, James ME, Charon NW, Manson MD, Norris SJ, Li C, Liu J. 2013. Cryoelectron tomography reveals the sequential assembly of bacterial flagella in *Borrelia burgdorferi*. Proc Natl Acad Sci 110:14390–14395.
130. Jones CJ, Macnab RM, Okino H, Aizawa S. 1990. Stoichiometric analysis of the flagellar hook-(basal-body) complex of *Salmonella typhimurium*. J Mol Biol 212:377–87.
131. Minamino T, Yamaguchi S, Macnab RM. 2000. Interaction between *fliE* and

flgB, a proximal rod component of the flagellar basal body of *Salmonella*. J Bacteriol 182:3029–3036.

132. Saijo-Hamano Y, Uchida N, Namba K, Oosawa K. 2004. In vitro characterization of FlgB, FlgC, FlgF, FlgG, and FliE, flagellar basal body proteins of *Salmonella*. J Mol Biol 339:423–35.
133. Osorio-Valeriano M, de la Mora J, Camarena L, Dreyfus G. 2016. Biochemical Characterization of the Flagellar Rod Components of *Rhodobacter sphaeroides*: Properties and Interactions. J Bacteriol 198:544–552.
134. Kubori T, Shimamoto N, Yamaguchi S, Namba K, Aizawa S. 1992. Morphological pathway of flagellar assembly in *Salmonella typhimurium*. J Mol Biol 226:433–46.
135. Cohen EJ, Ferreira JL, Ladinsky MS, Beeby M, Hughes KT. 2017. Nanoscale-length control of the flagellar driveshaft requires hitting the tethered outer membrane. Science 356:197–200.
136. Chevance FF V, Takahashi N, Karlinsey JE, Gnerer J, Hirano T, Samudrala R, Aizawa SI, Hughes KT. 2007. The mechanism of outer membrane penetration by the eubacterial flagellum and implications for spirochete evolution. Genes Dev 21:2326–2335.
137. Cohen EJ, Hughes KT. 2014. Rod-to-Hook transition for extracellular flagellum assembly is catalyzed by the L-ring-dependent rod scaffold removal. J Bacteriol 196:2387–2395.
138. Fujii T, Kato T, Namba K. 2009. Specific Arrangement of α-Helical Coiled Coils in the Core Domain of the Bacterial Flagellar Hook for the Universal Joint Function. Structure 17:1485–1493.
139. Brown MT, Steel BC, Silvestrin C, Wilkinson DA, Delalez NJ, Lumb CN, Obara B, Armitage JP, Berry RM. 2012. Flagellar hook flexibility is essential for bundle formation in swimming *Escherichia coli* cells. J Bacteriol 194:3495–3501.
140. Samatey FA, Matsunami H, Imada K, Nagashima S, Shaikh TR, Thomas DR, Chen JZ, Derosier DJ, Kitao A, Namba K. 2004. Structure of the bacterial flagellar hook and implication for the molecular universal joint mechanism. Nature 431:1062–8.
141. Hirano T, Yamaguchi S, Oosawa K, Aizawa S. 1994. Roles of FliK and FlhB in

determination of flagellar hook length in *Salmonella typhimurium*. J Bacteriol 176:5439–5449.

142. Erhardt M, Singer HM, Wee DH, Keener JP, Hughes KT. 2011. An infrequent molecular ruler controls flagellar hook length in *Salmonella enterica*. EMBO J 30:2948–61.
143. Homma M, Iino T. 1985. Locations of hook-associated proteins in flagellar structures of *Salmonella typhimurium*. J Bacteriol 162:183–189.
144. Ikeda T, Homma M, Iino T, Asakura S, Kamiya R. 1987. Localization and stoichiometry of hook-associated proteins within *Salmonella typhimurium* flagella. J Bacteriol 169:1168–73.
145. Paradis G, Chevance FF V., Liou W, Renault TT, Hughes KT, Rainville S, Erhardt M. 2017. Variability in bacterial flagella re-growth patterns after breakage. Sci Rep 7:1282.
146. Homma M, Iino T. 1985. Excretion of unassembled hook-associated proteins by *Salmonella typhimurium*. J Bacteriol 164:1370–1372.
147. Imada K, Vonderviszt F, Furukawa Y, Oosawa K, Namba K. 1998. Assembly characteristics of flagellar cap protein HAP2 of *Salmonella*: decamer and pentamer in the pH-sensitive equilibrium. J Mol Biol 277:883–91.
148. Yonekura K, Maki S, Morgan DG, DeRosier DJ, Vonderviszt F, Imada K, Namba K. 2000. The Bacterial Flagellar Cap as the Rotary Promoter of Flagellin Self-Assembly. Science 290:2148–2152.
149. Song WS, Cho SY, Hong HJ, Park SC, Yoon S. 2017. Self-Oligomerizing Structure of the Flagellar Cap Protein FliD and Its Implication in Filament Assembly. J Mol Biol 1–11.
150. Mimori Y, Yamashita I, Murata K, Fujiyoshi Y, Yonekura K, Toyoshima C, Namba K. 1995. The structure of the R-type straight flagellar filament of *Salmonella* at 9 Å resolution by electron cryomicroscopy. J Mol Biol 249:69–87.
151. Morgan DG, Owen C, Melanson L a, DeRosier DJ. 1995. Structure of bacterial flagellar filaments at 11 Å resolution: packing of the alpha-helices. J Mol Biol 249:88–110.
152. Mimori-Kiyosue Y, Vonderviszt F, Yamashita I, Fujiyoshi Y, Namba K. 1996. Direct interaction of flagellin termini essential for polymorphic ability of flagellar

filament. Proc Natl Acad Sci U S A 93:15108–15113.

153. Trachtenberg S, Fishelov D, Ben-Artzi M. 2003. Bacterial flagellar microhydrodynamics: Laminar flow over complex flagellar filaments, analog archimedean screws and cylinders, and its perturbations. Biophys J 85:1345–57.
154. Macnab RM, Ornston MK. 1977. Normal-to-curly flagellar transitions and their role in bacterial tumbling. Stabilization of an alternative quaternary structure by mechanical force. J Mol Biol 112:1–30.
155. Turner L, Ryu WS, Berg HC. 2000. Real-time imaging of fluorescent flagellar filaments. J Bacteriol 182:2793–2801.
156. Yamashita I, Hasegawa K, Suzuki H, Vonderviszt F, Mimori-Kiyosue Y, Namba K. 1998. Structure and switching of bacterial flagellar filaments studied by X-ray fiber diffraction. Nat Struct Biol 5:125–32.
157. Samatey FA, Imada K, Nagashima S, Vonderviszt F, Kumasaka T, Yamamoto M, Namba K. 2001. Structure of the bacterial flagellar protofilament and implications for a switch for supercoiling. Nature 410:331–7.
158. Maki-Yonekura S, Yonekura K, Namba K. 2010. Conformational change of flagellin for polymorphic supercoiling of the flagellar filament. Nat Struct Mol Biol 17:417–422.
159. Homma M, Komeda Y, Iino T, Macnab RM. 1987. The flaFIX gene product of *Salmonella typhimurium* is a flagellar basal body component with a signal peptide for export. J Bacteriol 169:1493–8.
160. Ibuki T, Imada K, Minamino T, Kato T, Miyata T, Namba K. 2011. Common architecture of the flagellar type III protein export apparatus and F- and V-type ATPases. Nat Struct Mol Biol 18:277–282.
161. Fraser GM, Bennett JCQ, Hughes C. 1999. Substrate-specific binding of hook-associated proteins by FlgN and FliT, putative chaperones for flagellum assembly. Mol Microbiol 32:569–580.
162. Auvray F, Thomas J, Fraser GM, Hughes C. 2001. Flagellin polymerisation control by a cytosolic export chaperone. J Mol Biol 308:221–229.
163. Thomas J, Stafford GP, Hughes C. 2004. Docking of cytosolic chaperone-substrate complexes at the membrane ATPase during flagellar type III protein export. Proc Natl Acad Sci U S A 101:3945–3950.

164. Imada K, Minamino T, Kinoshita M, Furukawa Y, Namba K. 2010. Structural insight into the regulatory mechanisms of interactions of the flagellar type III chaperone FliT with its binding partners. Proc Natl Acad Sci U S A 107:8812–8817.
165. Evans LDB, Poulter S, Terentjev EM, Hughes C, Fraser GM. 2013. A chain mechanism for flagellum growth. Nature 504:287–90.
166. Abrusci P, Vergara-Irigaray M, Johnson S, Beeby MD, Hendrixson DR, Roversi P, Friede ME, Deane JE, Jensen GJ, Tang CM, Lea SM. 2013. Architecture of the major component of the type III secretion system export apparatus. Nat Struct Mol Biol 20:99–104.
167. Bange G, Kümmerer N, Engel C, Bozkurt G, Wild K, Sinning I. 2010. FlhA provides the adaptor for coordinated delivery of late flagella building blocks to the type III secretion system. Proc Natl Acad Sci U S A 107:11295–11300.
168. Kinoshita M, Hara N, Imada K, Namba K, Minamino T. 2013. Interactions of bacterial flagellar chaperone-substrate complexes with FlhA contribute to co-ordinating assembly of the flagellar filament. Mol Microbiol 90:1249–1261.
169. Hara N, Namba K, Minamino T. 2011. Genetic characterization of conserved charged residues in the bacterial flagellar type III export protein FlhA. PLoS One 6.
170. Minamino T, Morimoto Y V, Hara N, Namba K. 2011. An energy transduction mechanism used in bacterial flagellar type III protein export. Nat Commun 2:475.
171. Galán JE. 2008. Energizing type III secretion machines: what is the fuel? Nat Struct Mol Biol 15:127–128.
172. Turner L, Stern AS, Berg HC. 2012. Growth of flagellar filaments of *Escherichia coli* is independent of filament length. J Bacteriol 194:2437–42.
173. O'Shea EK, Klemm JD, Kim PS, Alber T. 1991. X-ray structure of the GCN4 leucine zipper, a two-stranded, parallel coiled coil. Science 254:539–44.
174. Fezoui Y, Connolly P, Osterhout JJ. 1997. Solution structure of αtα, a helical hairpin peptide of de novo design. Protein Sci 6:1869–1877.
175. Komoriya K, Shibano N, Higano T, Azuma N, Yamaguchi S, Aizawa S-II. 1999. Flagellar proteins and type III-exported virulence factors are the predominant proteins secreted into the culture media of *Salmonella typhimurium*. Mol

Microbiol 34:767–79.

176. Furukawa Y, Imada K, Vonderviszt F, Matsunami H, Sano KI, Kutsukake K, Namba K. 2002. Interactions between bacterial flagellar axial proteins in their monomeric state in solution. J Mol Biol 318:889–900.
177. Portaliou AG, Tsolis KC, Loos MS, Zorzini V, Economou A. 2016. Type III Secretion: Building and Operating a Remarkable Nanomachine. Trends Biochem Sci 41:175–189.
178. Kubori T, Sukhan A, Aizawa SI, Galán JE. 2000. Molecular characterization and assembly of the needle complex of the *Salmonella typhimurium* type III protein secretion system. Proc Natl Acad Sci U S A 97:10225–30.
179. Journet L, Agrain C, Broz P, Cornelis GR. 2003. The Needle Length of Bacterial Injectisomes Is Determined by a Molecular Ruler. Science 302:1757–1760.
180. Minamino T, Saijo-Hamano Y, Furukawa Y, González-Pedrajo B, Macnab RM, Namba K. 2004. Domain organization and function of *Salmonella* FliK, a flagellar hook-length control protein. J Mol Biol 341:491–502.
181. Thomas NA, Finlay BB. 2004. Pathogens: Bacterial needles ruled to length and specificity. Curr Biol 14:192–194.
182. Diepold A, Wagner S. 2014. Assembly of the bacterial type III secretion machinery. FEMS Microbiol Rev 38:802–822.
183. Hueck CJ. 1998. Type III protein secretion systems in bacterial pathogens of animals and plants. Microbiol Mol Biol Rev 62:379–433.
184. Zarivach R, Deng W, Vuckovic M, Felise HB, Nguyen H V, Miller SI, Finlay BB, Strynadka NCJ. 2008. Structural analysis of the essential self-cleaving type III secretion proteins EscU and SpaS. Nature 453:124–7.
185. Weise CF, Login FH, Ho O, Gröbner G, Wolf-Watz H, Wolf-Watz M. 2014. Negatively Charged Lipid Membranes Promote a Disorder-Order Transition in the *Yersinia* YscU Protein. Biophys J 107:1950–1961.
186. Zilkenat S, Franz-Wachtel M, Stierhof Y-D, Galán JE, Macek B, Wagner S. 2016. Determination of the Stoichiometry of the Complete Bacterial Type III Secretion Needle Complex Using a Combined Quantitative Proteomic Approach. Mol Cell Proteomics 15:1598–1609.
187. Van Arnam JS, McMurry JL, Kihara M, Macnab RM. 2004. Analysis of an

Engineered *Salmonella* Flagellar Fusion Protein, FliR-FlhB. J Bacteriol 186:2495–2498.

188. McMurry JL, Minamino T, Furukawa Y, Francis JW, Hill SA, Helms KA, Namba K. 2015. Weak interactions between *Salmonella enterica* FlhB and other flagellar export apparatus proteins govern type III secretion dynamics. PLoS One 10:1–14.
189. Ferris HU, Furukawa Y, Minamino T, Kroetz MB, Kihara M, Namba K, Macnab RM. 2005. FlhB regulates ordered export of flagellar components via autocleavage mechanism. J Biol Chem 280:41236–41242.
190. Minamino T, Macnab RM. 2000. Domain structure of *Salmonella* FlhB, a flagellar export component responsible for substrate specificity switching. J Bacteriol 182:4906–14.
191. Morris DP, Roush ED, Thompson JW, Moseley MA, Murphy JW, McMurry JL. 2010. Kinetic characterization of *Salmonella* FliK-FlhB interactions demonstrates complexity of the Type III secretion substrate-specificity switch. Biochemistry 49:6386–93.
192. Meshcheryakov VA, Kitao A, Matsunami H, Samatey FA. 2013. Inhibition of a type III secretion system by the deletion of a short loop in one of its membrane proteins. Acta Crystallogr Sect D Biol Crystallogr 69:812–820.
193. Fraser GM, Hirano T, Ferris HU, Devgan LL, Kihara M, Macnab RM. 2003. Substrate specificity of type III flagellar protein export in *Salmonella* is controlled by subdomain interactions in FlhB. Mol Microbiol 48:1043–57.
194. Lavander M, Sundberg L, Edqvist PJ, Lloyd SA, Wolf-Watz H, Forsberg Å. 2002. Proteolytic cleavage of the FlhB homologue YscU of *Yersinia pseudotuberculosis* is essential for bacterial survival but not for type III secretion. J Bacteriol 184:4500–4509.
195. Deane JE, Graham SC, Mitchell EP, Flot D, Johnson S, Lea SM. 2008. Crystal structure of Spa40, the specificity switch for the *Shigella flexneri* type III secretion system. Mol Microbiol 69:267–76.
196. Morello JE, Collmer A. 2009. *Pseudomonas syringae* HrpP is a type III secretion substrate specificity switch domain protein that is translocated into plant cells but functions atypically for a substrate-switching protein. J Bacteriol 191:3120–3131.

197. Thomassin J-L, He X, Thomas NA. 2011. Role of EscU auto-cleavage in promoting type III effector translocation into host cells by enteropathogenic *Escherichia coli*. BMC Microbiol 11:205.
198. Frost S, Ho O, Login FH, Weise CF, Wolf-Watz H, Wolf-Watz M. 2012. Autoproteolysis and intramolecular dissociation of *Yersinia* YscU precedes secretion of its C-terminal polypeptide YscU(CC). PLoS One 7:e49349.
199. Login FH, Wolf-Watz H. 2015. YscU/FlhB of *Yersinia pseudotuberculosis* Harbors a C-terminal Type III Secretion Signal. J Biol Chem 290:26282–26291.
200. Minamino T, Doi H, Kutsukate K. 1999. Substrate specificity switching of the flagellum-specific export apparatus during flagellar morphogenesis in *Salmonella typhimurium*. Biosci Biotechnol Biochem 63:1301–1303.
201. Agrain C, Callebaut I, Journet L, Sorg I, Paroz C, Mota LJ, Cornelis GR. 2005. Characterization of a Type III secretion substrate specificity switch (T3S4) domain in YscP from *Yersinia enterocolitica*. Mol Microbiol 56:54–67.
202. Minamino T, Ferris HU, Moriya N, Kihara M, Namba K. 2006. Two parts of the T3S4 domain of the hook-length control protein FliK are essential for the substrate specificity switching of the flagellar type III export apparatus. J Mol Biol 362:1148–58.
203. Mizuno S, Amida H, Kobayashi N, Aizawa S-I, Tate S-I. 2011. The NMR structure of FliK, the trigger for the switch of substrate specificity in the flagellar type III secretion apparatus. J Mol Biol 409:558–73.
204. Shibata S, Takahashi N, Chevance FF V, Karlinsey JE, Hughes KT, Aizawa S-I. 2007. FliK regulates flagellar hook length as an internal ruler. Mol Microbiol 64:1404–15.
205. Kodera N, Uchida K, Ando T, Aizawa SI. 2015. Two-ball structure of the flagellar hook-length control protein FliK as revealed by high-speed atomic force microscopy. J Mol Biol 427:406–414.
206. Wagner S, Stenta M, Metzger LC, Dal Peraro M, Cornelis GR. 2010. Length control of the injectisome needle requires only one molecule of Yop secretion protein P (YscP). Proc Natl Acad Sci U S A 107:13860–5.
207. Agrain C, Sorg I, Paroz C, Cornelis GR. 2005. Secretion of YscP from *Yersinia enterocolitica* is essential to control the length of the injectisome needle but not

to change the type III secretion substrate specificity. Mol Microbiol 57:1415–1427.

208. Stainier I, Bleves S, Josenhans C, Karmani L, Kerbourch C, Lambermont I, Tötemeyer S, Boyd A, Cornelis GR. 2000. YscP, a *Yersinia* protein required for Yop secretion that is surface exposed, and released in low Ca2+. Mol Microbiol 37:1005–1018.
209. Moriya N, Minamino T, Hughes KT, Macnab RM, Namba K. 2006. The type III flagellar export specificity switch is dependent on FliK ruler and a molecular clock. J Mol Biol 359:466–77.
210. Hirano T, Shibata S, Ohnishi K, Tani T, Aizawa S-I. 2005. N-terminal signal region of FliK is dispensable for length control of the flagellar hook. Mol Microbiol 56:346–60.
211. Lorenz C, Schulz S, Wolsch T, Rossier O, Bonas U, Büttner D. 2008. HpaC controls substrate specificity of the *Xanthomonas* type III secretion system. PLoS Pathog 4.
212. Hirano T, Mizuno S, Aizawa S-I, Hughes KT. 2009. Mutations in *flk*, *flgG*, *flhA*, and *flhE* that affect the flagellar type III secretion specificity switch in *Salmonella enterica*. J Bacteriol 191:3938–49.
213. Aldridge P, Karlinsey JE, Becker E, Chevance FF V, Hughes KT. 2006. Flk prevents premature secretion of the anti-σ factor FlgM into the periplasm. Mol Microbiol 60:630–642.
214. Karlinsey JE, Tsui HCT, Winkler ME, Hughes KT. 1998. Flk couples *flgM* translation to flagellar ring assembly in *Salmonella typhimurium*. J Bacteriol 180:5384–5397.
215. Stafford GP, Hughes C. 2007. *Salmonella typhimurium flhE*, a conserved flagellar regulon gene required for swarming. Microbiology 153:541–547.
216. Lee J, Harshey RM. 2012. Loss of FlhE in the flagellar type III secretion system allows proton influx into *Salmonella* and *Escherichia coli*. Mol Microbiol 84:550–565.
217. Lee J, Monzingo AF, Keatinge-Clay AT, Harshey RM. 2015. Structure of *Salmonella* FlhE, conserved member of a flagellar type III secretion operon. J Mol Biol 427:1254–1262.
218. Makishima S, Komoriya K, Yamaguchi S, Aizawa S-I. 2001. Length of the

flagellar hook and the capacity of the type III export apparatus. Science 291:2411–2413.

219. Shaikh TR, Thomas DR, Chen JZ, Samatey FA, Matsunami H, Imada K, Namba K, Derosier DJ. 2005. A partial atomic structure for the flagellar hook of *Salmonella typhimurium*. Proc Natl Acad Sci U S A 102:1023–1028.
220. Stern AS, Berg HC. 2013. Single-file diffusion of flagellin in flagellar filaments. Biophys J 105:182–184.
221. Blank K, Hensel M, Gerlach RG. 2011. Rapid and highly efficient method for scarless mutagenesis within the *Salmonella enterica* chromosome. PLoS One 6:e15763.
222. Meshcheryakov VA, Yoon Y-H, Samatey FA. 2011. Purification, crystallization and preliminary X-ray crystallographic analysis of the C-terminal cytoplasmic domain of FlhB from *Aquifex aeolicus*. Acta Crystallogr Sect F Struct Biol Cryst Commun 67:280–282.
223. Meshcheryakov V a, Samatey F a. 2011. Purification, crystallization and preliminary X-ray crystallographic analysis of the C-terminal cytoplasmic domain of FlhB from *Salmonella typhimurium*. Acta Crystallogr Sect F Struct Biol Cryst Commun 67:808–11.
224. Battye TGG, Kontogiannis L, Johnson O, Powell HR, Leslie AGW. 2011. iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. Acta Crystallogr D Biol Crystallogr 67:271–81.
225. Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, Keegan RM, Krissinel EB, Leslie AGW, McCoy A, McNicholas SJ, Murshudov GN, Pannu NS, Potterton E a, Powell HR, Read RJ, Vagin A, Wilson KS. 2011. Overview of the CCP4 suite and current developments. Acta Crystallogr D Biol Crystallogr 67:235–42.
226. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. 2007. Phaser crystallographic software. J Appl Crystallogr 40:658–674.
227. Murshudov GN, Vagin a a, Dodson EJ. 1997. Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr D Biol Crystallogr 53:240–55.
228. Vranken WF, Boucher W, Stevens TJ, Fogh RH, Pajon A, Llinas M, Ulrich EL, Markley JL, Ionides J, Laue ED. 2005. The CCPN data model for NMR

spectroscopy: Development of a software pipeline. Proteins Struct Funct Genet 59:687–696.

229. Williamson MP. 2013. Using chemical shift perturbation to characterise ligand binding. Prog Nucl Magn Reson Spectrosc 73:1–16.

230. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. 2000. The protein data bank. Nucleic Acids Res 28:235–242.

231. Bateman A, Martin MJ, O'Donovan C, Magrane M, Apweiler R, Alpi E, Antunes R, Arganiska J, Bely B, Bingley M, Bonilla C, Britto R, Bursteinas B, Chavali G, Cibrian-Uhalte E, Da Silva A, De Giorgi M, Dogan T, Fazzini F, Gane P, Castro LG, Garmiri P, Hatton-Ellis E, Hieta R, Huntley R, Legge D, Liu W, Luo J, Macdougall A, Mutowo P, Nightingale A, Orchard S, Pichler K, Poggioli D, Pundir S, Pureza L, Qi G, Rosanoff S, Saidi R, Sawford T, Shypitsyna A, Turner E, Volynkin V, Wardell T, Watkins X, Zellner H, Cowley A, Figueira L, Li W, McWilliam H, Lopez R, Xenarios I, Bougueleret L, Bridge A, Poux S, Redaschi N, Aimo L, Argoud-Puy G, Auchincloss A, Axelsen K, Bansal P, Baratin D, Blatter MC, Boeckmann B, Bolleman J, Boutet E, Breuza L, Casal-Casas C, De Castro E, Coudert E, Cuche B, Doche M, Dornevil D, Duvaud S, Estreicher A, Famiglietti L, Feuermann M, Gasteiger E, Gehant S, Gerritsen V, Gos A, Gruaz-Gumowski N, Hinz U, Hulo C, Jungo F, Keller G, Lara V, Lemercier P, Lieberherr D, Lombardot T, Martin X, Masson P, Morgat A, Neto T, Nouspikel N, Paesano S, Pedruzzi I, Pilbout S, Pozzato M, Pruess M, Rivoire C, Roechert B, Schneider M, Sigrist C, Sonesson K, Staehli S, Stutz A, Sundaram S, Tognolli M, Verbregue L, Veuthey AL, Wu CH, Arighi CN, Arminski L, Chen C, Chen Y, Garavelli JS, Huang H, Laiho K, McGarvey P, Natale DA, Suzek BE, Vinayaka CR, Wang Q, Wang Y, Yeh LS, Yerramalla MS, Zhang J. 2015. UniProt: A hub for protein information. Nucleic Acids Res 43:D204–D212.

232. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Söding J, Thompson JD, Higgins DG. 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol Syst Biol 7:539.

233. Greenfield NJ. 2007. Using circular dichroism collected as a function of

temperature to determine the thermodynamics of protein unfolding and binding interactions. Nat Protoc 1:2527–2535.

234. Kemmer G, Keller S. 2010. Nonlinear least-squares data fitting in Excel spreadsheets. Nat Protoc 5:267–281.
235. Koronakis V, Hughes C, Koronakis E. 1991. Energetically distinct early and late stages of HlyB/HlyD-dependent secretion across both *Escherichia coli* membranes. EMBO J 10:3263–72.
236. Mahon CS, O'Donoghue AJ, Goetz DH, Murray PG, Craik CS, Tuohy MG. 2009. Characterization of a multimeric, eukaryotic prolyl aminopeptidase: An inducible and highly specific intracellular peptidase from the non-pathogenic fungus *Talaromyces emersonii*. Microbiology 155:3673–3682.
237. Meshcheryakov VA, Samatey FA. 2011. Purification, crystallization and preliminary X-ray crystallographic analysis of the C-terminal cytoplasmic domain of FlhB from *Salmonella typhimurium*. Acta Crystallogr Sect F Struct Biol Cryst Commun 67:808–811.
238. Hirano T, Minamino T, Namba K, Macnab RM. 2003. Substrate specificity classes and the recognition signal for *Salmonella* type III flagellar export. J Bacteriol 185:2485–2492.
239. Siegele D a, Hu JC. 1997. Gene expression from plasmids containing the *araBAD* promoter at subsaturating inducer concentrations represents mixed populations. Proc Natl Acad Sci U S A 94:8168–8172.
240. Uchida K, Dono K, Aizawa S-I. 2016. Identification of the Key Sequence in the FliK C-Terminal Domain for Substrate Specificity Switching in the Flagellar Protein Secretion. J Bacteriol 198:410–415.
241. Meshcheryakov VA, Barker CS, Kostyukova AS, Samatey FA. 2013. Function of FlhB, a Membrane Protein Implicated in the Bacterial Flagellar Type III Secretion System. PLoS One 8:1–13.
242. Kutsukake K, Minamino T, Yokoseki T. 1994. Isolation and characterization of FliK-independent flagellation mutants from *Salmonella typhimurium*. J Bacteriol 176:7625–7629.
243. Williams AW, Yamaguchi S, Togashi F, Aizawa S, Kawagishi I, Macnab RM. 1996. Mutations in *fliK* and *flhB* affecting flagellar hook and filament assembly in *Salmonella typhimurium*. J Bacteriol 178:2960–2970.

244. Edqvist PJ, Olsson J, Lavander M, Forsberg Å, Wolf-watz H, Scott A, Sundberg L, Lloyd SA. 2003. YscP and YscU Regulate Substrate Specificity of the *Yersinia* Type III Secretion System. J Bacteriol 185:2259–2266.
245. Ryu J, Hartin R. 1990. Quick transformation in *Salmonella typhimurium* LT2. Biotechniques 8:43–44.
246. Datsenko KA, Wanner BL. 2000. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. Proc Natl Acad Sci U S A 97:6640–5.
247. Miroux B, Walker JE. 1996. Over-production of proteins in *Escherichia coli*: mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. J Mol Biol 260:289–98.
248. Guzman LM, Belin D, Carson MJ, Beckwith J. 1995. Tight regulation, modulation, and high-level expression by vectors containing the arabinose P_{BAD} promoter. J Bacteriol 177:4121–30.

Appendix A

A1.1 Reagents, buffers, and media

All buffers and media were sterilised by filtration or autoclaving.

Name	Composition
Luria-Bertani (LB)-Lennox broth	1% w/v tryptone 0.5% w/v NaCl 0.5% w/v yeast extract
LB-Lennox agar	1% w/v tryptone 0.5% w/v NaCl 0.5% w/v yeast extract 1.5% w/v agar
Motility agar	0.25% w/v Bacto-agar 1% w/v tryptone 0.5% w/v NaCl
2x Yeast Tryptone (2xYT) broth	1.6% w/v tryptone 1.0% w/v yeast extract 0.5% w/v NaCl
Super optimal broth with catabolite repression (SOC)	2% w/v tryptone 0.5% w/v yeast extract 10 mM NaCl 2.5 mM KCl 10 mM $MgCl_2$ 10 mM $MgSO_4$ 20 mM D-glucose
M9 minimal medium	0.5% w/v D-glucose (Carbon-13, all 6 carbons labeled ($^{13}C_6$), 99% atom ^{13}C, as needed) 0.1% w/v NH_4Cl (Nitrogen-15 labeled, 99% atom ^{15}N, as needed) 42.3 mM Na_2HPO_4 22 mM KH_2PO_4 8.6 mM NaCl 5 mM $MgSO_4$ 0.1 mM $CaCl_2$ 4 μM $ZnSO_4$ 1 μM $MnCl_2$ 0.7 μM H_3BO_3 0.7 μM $CuSO_4$ 2 μM $FeCl_3$ Sterile water to 1 L
M9 minimal medium with D_2O and deuterated glucose	Same as above, but sterile water was replaced with 100% D_2O (Sigma Aldrich) as needed. D-Glucose was

	replaced with 0.5% w/v D-glucose, $^{13}C_6$-1,2,3,4,5,6,6-deuterium$_7$, 99% atom ^{13}C, 97% deuterium (Sigma Aldrich) as needed.
Buffer A	20 mM Tris-HCl, pH 8.0 1 M NaCl 20% v/v glycerol 2 mM EDTA
Buffer B	50 mM Tris-HCl pH 7.4 150 mM NaCl
Buffer C	20 mM sodium phosphate buffer, pH 7.4 100 mM NaCl
Buffer D	20 mM Tris-HCl, pH 7.4 500 mM NaCl 20 mM imidazole
Buffer E	20 mM Tris-HCl, pH 7.4 500 mM NaCl 300 mM imidazole
Buffer F	20 mM Tris-HCl, pH 8.0 150 mM NaCl 20 mM imidazole (300 mM for co-elution)
Buffer G	20 mM Tris-HCl, pH 7.4 150 mM NaCl
SDS loading buffer, 2x	100 mM Tris-Hcl, pH 8.8 4% w/v sodium dodecyl sulfate (SDS) 20% v/v glycerol 0.2% w/v Bromophenol Blue 140 mM β-mercaptoethanol (βME)
SDS running buffer, 1x	192 mM glycine 25 mM Tris-HCl, pH 8.3 0.1% w/v SDS
Transfer buffer, 1x	10 mM 3-(Cyclohexylamino)-1-propanesulfonic acid (CAPS), pH 11 10% v/v methanol
Tris-Acetate-EDTA (TAE) buffer, 1x	40 mM Tris-HCl, pH 8.5 20 mM acetic acid 1 mM EDTA
Phosphate buffered saline (PBS), 10x	1.37 M NaCl 27 mM KCl 100 mM Na_2HPO_4 18 mM KH_2PO_4 water to 1 L, pH 7.4
Phosphate buffered saline with 0.05% Triton X-100 (PBST), 1x	1x PBS 0.05% Triton X-100

A1.2 Oligonucleotides

Name	Description & Restriction Site	Destination Plasmid	Sequence (5'→3')	Length
T7For	Sequencing primer	pTXBI and pET15b	TAATACGACTCACTATAGGG	20
T7Rev	Sequencing primer	pTXBI and pET15b	GCTAGTTATTGCTCAGCGGTGG	22
pTXBI inteinRev	Sequencing primer	pTXBI	GATGTCGGCGATGCGTAC	18
pBADFor	Sequencing primer	pBAD18amp	GATTATTTGCACGGCGTCACACTTTGCTATGC	32
pBADRev	Sequencing primer	pBAD18amp	CTGCGTTCTGATTTAATCTGTATCAGGCTG	30
FliKFor	NheI, Ribosome binding site (RBS) from pET15b	pBAD18amp	GCTATTGCTAGCAGGAGAGCCCAAATGATCACCCTGCCCCAACTGATC	48
FliK270For	NheI, RBS (from pET15b)	pBAD18amp	GCTATTGCTAGCAGGAGAGCCCAAATGGAATGGCAGCAAACGTTCAGT	48
FliK405 StopRev	XbaI	pBAD18amp	GCATACTCTAGATTAGGCGAAGATATCCACTGCGCC	36
FliKFor	NdeI	pET15b	GCTATTCATATGATCACCCTGCCCCAACTG	30
FliK270For	NdeI	pET15b	GCTATTCATATGGAATGGCAGCAAACGTTC	30
FliK405 StopRev	BamHI	pET15b	GCATACGGATCCTTAGGCGAAGATATCCACTGC	33
FliKE297GE298GFor	Overlap extension mutagenesis primer	pBAD18amp	CGATTGCATCCGGGCGGCTTAGGTCAGGTG	30
FliKE297GE298GRev	Overlap extension mutagenesis primer	pBAD18amp	CACCTGACCTAAGCCGCCCGGATGCAATCG	30
FliKD309GD310GFor	Overlap extension mutagenesis primer	pBAD18amp	TCGCTCAAGCTGGGCGGCAATCAGGCGCAG	30
FliKD309GD310GRev	Overlap extension mutagenesis primer	pBAD18amp	CTGCGCCTGATTGCCGCCCAGCTTGAGCGA	30
FliK717For	For replacement of FliK with Kanamycin^{R}	pWRG717 (template)	AAACCCGAATGATCACCCTGCCCCAACTGATCACCACCGATACCGAGGGTTTTCCCAGTCACGAC	65
FliK717Rev	For replacement of FliK with Kanamycin^{R}	pWRG717 (template)	CAAAGCGTGGAAAAGACGCGGATAATCATGCTACCTCTGGCGTTATGCTTCCGGCTCGTATGTTG	65
FlhBFor211	NdeI	pTXBI	GCTATTCATATGTTCCAGATCTTTAGCCACCTG	33

FlhBFor219	NdeI	pTXBI	GCTATTCATATGAAATTACGCATGT CGCGGCAG	33
FlhBFor270	NdeI	pTXBI	GCTATTCATATGCCGACGCACTATT CCGTGGCG	33
FlhBcc G_ARev	SapI	pTXBI	GGTGGTTGCTCTTCCGCACGCATC AGTATTCTTCTCGTT	39
FlhBcnRev	SapI	pTXBI	GGTGGTTGCTCTTCCGCAGTTAGT GACAATGACGTCCGC	39
FlhBcc W353TRev	SapI	pTXBI	GGTGGTTGCTCTTCCGCAGGTGCG TTTAAGCTGCCAGAC	39
FlhBc N269AFor	Overlap extension mutagenesis primer	pTXBI	GTCATTGTCACTGCGCCGACGCAC TATTCC	30
FlhBc N269ARev	Overlap extension mutagenesis primer	pTXBI	GGAATAGTGCGTCGGCGCAGTGAC AATGAC	30
FlhBc P238AFor	Overlap extension mutagenesis primer	pTXBI	AGCGAAGGCGATGCGCATGTTAAG GGC	27
FlhBc P238ARev	Overlap extension mutagenesis primer	pTXBI	GCCCTTAACATGCGCATCGCCTTC GCT	27
FlhB S274Ffor	Overlap extension mutagenesis primer	pTXBI	ACGCACTATTTTGTGGCGCTG	21
FlhB S274Frev	Overlap extension mutagenesis primer		CAGCGCCACAAAATAGTGCGT	21
FlhB R302Pfor	Overlap extension mutagenesis primer		CTGCGCATTCCGGAGATCGGC	21
FlhB R302Prev	Overlap extension mutagenesis primer		GCCGATCTCCGGAATGCGCAG	21
FlhBseqfor	Flanking primer		GGACGACCTGCTGGACAGT	19
FlhBseqrev	Flanking primer		CGAGGATAAAAGCAGGTAGC	20
FlgC 717For	For replacement of FlgC with Kanamycin^{R}	pWRG717 (template)	GGACAGCCTTATCGCGCCAAACAG GTGGTTTTTCAGGTGGACGCCAGG GTTTTCCCAGTCACGAC	65
FlgC 717Rev	For replacement of FlgC with Kanamycin^{R}	pWRG717 (template)	AATCACGCTGGCGACCTTTACCCC GCCAGTGGCTTGACCCGGCGCTGC TTCCGGCTCGTATGTTG	65
FlgC KnockinFor	For replacement of Kanamycin^{R} cassette for scarless deletion of FlgC		AGGCATGATGAATGTGCTACAGGG AGGAAACTAATCCGTGTAAAGGAG GCGCGTATGTCTATTGCCGTAAATA TGAATGA	80
FlgC KnockinRe v	For replacement of Kanamycin^{R} cassette for scarless deletion of FlgC		TCATTCATATTTACGGCAATAGACA TACGCGCCTCCTTTACACGAATTAG TTTCCTCCCTGTAGCACATTCATCA TGCCT	80

FlgC SeqFor	Flanking primer to sequence FlgC and surrounding region		GTTTATTCCGGCGATAAC	18
FlgC SeqRev	Flanking primer to sequence FlgC and surrounding region		CTGCCGGTCGTCGTTTTGAC	20
FlgC F49AFor	Overlap extension mutagenesis primer		GCCAAACAGGTGGTTGCGCAGGTGGACGCCGCG	33
FlgC F49ARev	Overlap extension mutagenesis primer		CGCGGCGTCCACCTGCGCAACCACCTGTTTGGC	33
FlgC GRMswap For	Overlap extension mutagenesis primer		CGCGCCAAACAGGTGGTTTTTGCCAGTGAGTTAGCGCCGGGTCAAGCCACT	51
FlgC GRMswap Rev	Overlap extension mutagenesis primer		AGTGGCTTGACCCGGCGCTAACTCACTGGCAAAAACCACCTGTTTGGCGCG	51
FlgCFor	NheI, RBS (from pET15b	pBAD18amp	GCTATTGCTAGCAGGAGAGCCCAAGTGGCGCTGTTAAACATTTTT	45
FlgC StopRev	XbaI, with stop codon	pBAD18amp	GCATACTCTAGATTACTGGCCTAATGTCAGCGT	33
FlgE F39KFor	Overlap extension mutagenesis primer		GGCATCAAAAGCCGATATGTTCGCCGGT	28
FlgE F39KRev	Overlap extension mutagenesis primer		CCGGCGAACATATCGGCTTTTGATGCC	27
FlgE F43KFor	Overlap extension mutagenesis primer		GGCATCATTTGCCGATATGAAAGCCGGT	28
FlgE F43KRev	Overlap extension mutagenesis primer		CCGGCTTTCATATCGGCAAATGATGCC	27
FlgEF39K F43KFor	Overlap extension mutagenesis primer		GGCATCAAAAGCCGATATGAAAGCCGGT	28
FlgEF39K F43KRev	Overlap extension mutagenesis primer		CCGGCTTTCATATCGGCTTTTGATGCC	27
FlgE F39AFor	Overlap extension mutagenesis primer		GGCATCAGCGGCCGATATGTTCGCCGGT	28
FlgE F39ARev	Overlap extension mutagenesis primer		CCGGCGAACATATCGGCCGCTGATGCC	27
FlgE F43AFor	Overlap extension mutagenesis primer		GGCATCATTTGCCGATATGGCGGCCGGT	28
FlgE F43ARev	Overlap extension mutagenesis primer		CCGGCCGCCATATCGGCAAATGATGCC	27
FlgEF39A F43AFor	Overlap extension mutagenesis primer		GGCATCAGCGGCCGATATGGCGGCCGGT	28
FlgEF39A F43ARev	Overlap extension mutagenesis primer		CCGGCCGCCATATCGGCCGCTGATGCC	27

A1.3 Bacterial strains

Note: The source of strains includes whomever constructed the strain and the method used to construct the strain.

Strain	Characteristic	Source
***Salmonella enterica* serovar Typhimurium strains**		
JR501	For improved transformation of *E. coli* plasmids into *Salmonella*	(245)
SJW1103	Motile, phase-1 stable *H1-I H2-enx*$^{-,off}$ *vh2*$^{-}$	(58)
SJW1103*ΔflgD::Kan*^{R}	*flgD*$^{-}$, Kanamycin^{R}	L. Evans, (246)
SJW1103*ΔflgE::Kan*^{R}	*flgE*$^{-}$, Kanamycin^{R}	L. Evans, (246)
SJW1103*ΔfliK::FRT*	*fliK*$^{-}$, replaced with FLP recognition target (FRT) scar sequence	N. Greene, (246)
SJW1103*ΔflgC*	*flgC*$^{-}$	This study, (221)
SJW1103*flhB*$_{CT3xFLAG}$	*flhB* tagged with C-terminal 3xFLAG	O. Bryant, (221)
SJW1103*flhB*$_{CT3xFLAG}$ *N*$_{269}$*A*	*flhB* tagged with C-terminal 3xFLAG, $N_{269}A$ non-autocleavable variant	This study, (221)
SJW1103*flhB*$_{CT3xFLAG}$ *P*$_{238}$*A N*$_{269}$*A*	*flhB* tagged with C-terminal 3xFLAG, $P_{238}A$ $N_{269}A$ non-autocleavable variant	This study, (221)
SJW1103*flhB*$_{CT3xFLAG}$ *S*$_{274}$*F*	*flhB* tagged with C-terminal 3xFLAG, $S_{274}F$ autonomous switching mutant	This study, (221)
SJW1103*flhB*$_{CT3xFLAG}$ *R*$_{302}$*P*	*flhB* tagged with C-terminal 3xFLAG, $R_{302}P$ autonomous switching mutant	This study, (221)
SJW1103*flhB*$_{CT3xFLAG}$ *S*$_{274}$*F R*$_{302}$*P*	*flhB* tagged with C-terminal 3xFLAG, $S_{274}F$ $R_{302}P$ combined autonomous switching mutant	This study, (221)
SJW1103*ΔfliK::Kan*^{R} *flhB*$_{CT3xFLAG}$	*flhB* tagged with C-terminal 3xFLAG, *fliK*$^{-}$, Kanamycin^{R}	This study, (221)
SJW1103*ΔfliK::Kan*^{R} *flhB*$_{CT3xFLAG}$ *N*$_{269}$*A*	*flhB* tagged with C-terminal 3xFLAG, $N_{269}A$ non-autocleavable variant, *fliK*$^{-}$, Kanamycin^{R}	This study, (221)
SJW1103*ΔfliK::Kan*^{R} *flhB*$_{CT3xFLAG}$ *S*$_{274}$*F R*$_{302}$*P*	*flhB* tagged with C-terminal 3xFLAG, $S_{274}F$ $R_{302}P$ combined autonomous switching mutant, *fliK*$^{-}$, Kanamycin^{R}	This study, (221)
***Escherichia coli* strains**		
C41 (DE3)	F^{-} ompT gal dcm hsdS$_{B}$ (r$_{B}^{-}$ m$_{B}^{-}$)	(247)
XL-1	recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac [F′ proAB lacIqZΔM15 Tn10(T etr)]	Laboratory stock
DH10β	F– Φ80lacZΔM15 Δ(lacZYA- argF)U169 recA1 endA1 hsdR17 (rK–, mK+) phoA supE44 λ– thi-1 gyrA96 relA1	Laboratory Stock

A1.4 Plasmids

Note: All recombinant plasmids derived from parental vectors carry all features of the parental vector unless otherwise noted (e.g., all pTXBI constructs are C-terminally fused to the inteinCBD and expression is induced with IPTG). All genes originate from *Salmonella enterica* serovar Typhimurium strain SJW1103.

Plasmid	Characteristics	Source
pTXBI	6.7k base pairs, Ampicillin^R, protein expression under IPTG-inducible T7 promoter, C-terminal fusion with *Mxe* intein/chitin binding domain (inteinCBD) via a cysteine residue	New England Biolabs (NEB)
pTXBI-FlhB$_C$	FlhB$_C$ domain, residues 219-383, $G_{383}A$ for better cleavage from inteinCBD	(237)
pTXBI-FlhB$_{CN}$	Residues 211-269	This study
pTXBI-FlhB$_{CC}$	Residues 270-383, $G_{383}A$	This study
pTXBI-FlhB$_C$ $N_{269}A$	$G_{383}A$ $N_{269}A$ to prevent autocleavage at primary site	This study
pTXBI-FlhB$_C$ $P_{238}A$	FlhB$_C$ domain, amino acids 219-383, $G_{383}A$, $P_{238}A$ to prevent autocleavage at secondary site	This study
pTXBI-FlhB$_C$ $P_{238}A$ $N_{269}A$	$G_{383}A$ $P_{238}A$ and $N_{269}A$ to prevent autocleavage at primary and secondary site	This study
pTXBI-FlhB$_C$ $P_{287}E$	$G_{383}A$, $P_{287}E$ to abolish early subunit binding	This study
pTXBI-FlhB$_{C211}$	FlhB$_C$ domain, residues 211-383, $G_{383}A$	This study
pTXBI-FlhB$_{C211}$ $N_{269}A$	$G_{383}A$ $N_{269}A$ to prevent autocleavage at primary site	This study
pTXBI-FlhB$_{C211}$ $W_{353}T$	$G_{383}A$ $W_{353}T$ to mimic autonomous switching mutant W_{353}Stop	This study
pTXBI-FlhB$_{C211}$ $S_{274}F$	$G_{383}A$ $S_{274}F$ autonomous switcher	This study
pTXBI-FlhB$_{C211}$ $R_{302}P$	$G_{383}A$ $R_{302}P$ autonomous switcher	This study
pTXBI-FlhB$_{C211}$ $S_{274}F$ $R_{302}P$	$G_{383}A$ $S_{274}F$ $R_{302}P$ autonomous switcher	This study
pTXBI-FlhB$_{CCT3xFLAG}$	FlhB$_C$ domain, residues 211-383, C-terminal 3xFLAG	This study
pTXBI-HexaHis-3xFLAG-FlgB-FlgC	HexaHis-3xFLAG-FlgB and FlgC (C-terminal fusion to inteinCBD). This plasmid allowed for co-expression of both recombinant proteins. FlgB under the vector T7 promoter and FlgC	This study

	under the native *Salmonella* promoter found in FlgB.	
pBAD18amp	4.6k base pairs, Ampicillin^R, protein expression under the control of the P_{BAD} promoter of the *araBAD* operon	(248)
pBAD18amp-FlgE$_{mid3xFLAG}$	Residues 1-403 for expression in *S. enterica*, 3x mid-FLAG tag between residues 234 and 235.	L. Evans
pBAD18amp-FlgE$_{\Delta GRM}$	Residues 1-38, 44-403, 3x mid-FLAG tag	L. Evans
pBAD18amp-FlgE$_{F39K}$	Point mutation- phenylalanine 39 to lysine, 3x mid-FLAG tag	This study
pBAD18amp-FlgE$_{F43K}$	Point mutation- phenylalanine 43 to lysine, 3x mid-FLAG tag	This study
pBAD18amp-FlgE$_{F39KF43K}$	Point mutation- phenylalanine 39 and 43 to lysine, 3x mid-FLAG tag	This study
pBAD18amp-FlgE$_{F39A}$	Point mutation- phenylalanine 39 to alanine, 3x mid-FLAG tag	This study
pBAD18amp-FlgE$_{F43A}$	Point mutation- phenylalanine 43 to alanine, 3x mid-FLAG tag	This study
pBAD18amp-FlgE$_{F39AF43A}$	Point mutation- phenylalanine 39 and 43 to alanine, 3x mid-FLAG tag	This study
pBAD18amp-FlgC$_{wt}$	Residues 1-134 for expression in *S. enterica*	This study
pBAD18amp-FlgC$_{F49A}$	Point mutation- phenylalanine 49 to alanine	This study
pBAD18amp-FlgC$_{GRMswap}$	GRM swap with native DNA sequence from FlgB	This study
pBAD18amp-FliK$_{wt}$	Residues 1-405 for expression in *S. enterica*	This study
pBAD18amp-FliK$_{270}$	Residues 270-405 for expression in *S. enterica*	This study
pBAD18amp-FliK$_{EEGG}$	Residues 1-405 for expression in *S. enterica*, glutamic acid 297 and 298 to glycines	This study
pBAD18amp-FliK$_{DDGG}$	Residues 1-405 for expression in *S. enterica*, aspartic acid 309 and 310 to glycines	This study
pET15b	5.7k base pairs, Ampicillin^R, IPTG-inducible T7 promoter, N-terminal fusion of HexaHis-tag followed by thrombin cleavage site	Novagen
pET15b-FliK$_{wt}$	Residues 1-405	This study
pET15b-FliK$_{270}$	Residues 270-405	This study
pWRG717	Kanamycin^R with I-*Sce*I endonuclease recognition site	R. Gerlach (update on Blank, Hensel, and Gerlach, 2011)
pWRG730	Chloramphenicol^R. Lambda red recombinase under a heat-shock inducible promoter. I-*Sce*I endonuclease under a tetracycline inducible promoter.	R. Gerlach (update on Blank, Hensel, and Gerlach, 2011)

A1.5 Antibodies

Antibody	Host	Dilution	Source
Penta·His	Mouse, IgG, monoclonal	1:2000	Qiagen
Anti-FLAG M2	Mouse, IgG1, monoclonal	1:5000 (for $FlgE_{mid3xFLAG}$) 1:1000 (for $FlhB_{Ct3xFLAG}$)	Sigma Aldrich
Anti-chitin binding domain (anti-CBD)	Mouse, IgG, monoclonal	1:1000	NEB
IRDye 680RD anti-mouse IgG	Donkey	1:10000	Li-Cor
IRDye 800CW anti-rabbit IgG	Donkey	1:10000	Li-Cor
Anti-$FlhB_C$	Rabbit, polyclonal	1:2000	F. Samatey
Anti-FlgN	Rabbit, polyclonal	1:2000	Fraser Lab
Anti-FliC (Salmonella H Antiserum i)	Rabbit, polyclonal	1:5000	Difco (Becton, Dickinson, and Company)